


```

Db      520 ACCCCAGCCCC 531

RESULT 3
LOCUS   AV656559
DEFINITION AV656559 Bos taurus brain fetus Bos taurus cDNA clone E1BR04.9H02
5', mRNA sequence.
ACCESSION AV656559
VERSION   AV656559.1 GI:9924689
KEYWORDS EST.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovidae; Bovinae; Bos.
REFERENCE Takasuga,A., Hirotsune,S., Itoh,R., Uchizono,A., Suzuki,H., Aso,H.
AUTHORS   and Sugimoto,Y.
TITLE     Establishment of a high throughput EST sequencing system using
          poly(A) tail-removed cDNA libraries and determination of 36,000
          bovine ESTs
          Nucleic Acids Res. 29 (22), E108 (2001)
JOURNAL   21570554
MEDLINE   11713328
COMMENT   Contact: Yoshikazu Sugimoto
          Animal Genetics Division
          Shikawa Institute of Animal Genetics
          Odokura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
          Tel: 81-248-25-5641
          Fax: 81-248-25-5725
          Email: kazusugi@cocococ.nac.ne.jp
          Single pass sequencing.
          This clone was obtained from a polyA-deleted cDNA library.
FEATURES
SOURCE    Location/Qualifiers
          1..557
            /organism="Bos taurus"
            /mol_type="mRNA"
            /db_xref="taxon:9913"
            /clone="E1BR04.9H02"
            /tissue_type="brain"
            /dev_stage="fetus"
            /lab_host="DH10B"
            /clone_id="Bos taurus brain fetus"
            /note="Vector: pZ1; Site 1: SalI; Site 2: NotI; Poly A
          was deleted from a NotI site"
BASE COUNT      89 a      214 c      158 g      91 t      5 others
ORIGIN
Query Match      52.1%; Score 339; DB 9; Length 557;
Best Local Similarity 83.9%; Pred. No. 3.8e-60;
Matches 407; Conservative 0; Mismatches 69; Indels 9; Gaps 2;

QY      9 CGGCGCTGGAACCATGACACGGGCTATCCCGGAGAGACACCGGGCCCCCAAGGCCCA 68
Db      12 CGAGCTGGAACCATGACACGGGCTATCCCGGAGAGACACCGGGCCCCCAAGGCCCA 71

QY      69 GCAAGGCGCG-----GTGCCACACAGCCCTTCACTGGCGGCGCCCGACCCCGGCGCTC 122
Db      72 GCAAGGCGCGAGGCGACAGCCCGACAGCGCCCTGAGCGCTGGGGCTCCGAGACCCCACTTC 131

QY      123 GAGACCACTGATGTCGTGGTGTTCAGACACCTTCACTGAATGTCGTGGCTCGGCT 182
Db      132 GAGACCACTGATGTCGTGGTGTTCAGACACCTTCACTGAATGTCGTGGCTCGGCT 191

QY      183 TCCGCGCGCTGGCGCACTCCATCAAGGCGCGAGATCAAGAGTGGTGGTGAACCTCGAAG 242
Db      192 TCCGCGCGCTGGCGCACTCCATCAAGGCGCGAGATCAAGAGTGGTGGTGAACCTCGAAG 251

QY      243 CGGCGCGCGCTGGTTCCTCAAGGCGAGGCTCAACATCTGGCGCGGATGTCGAGCGC 302
Db      252 CAGCCCGCGCGCTGGTTCCTCAAGGCGAGGCTCAACATCTGGCGCGGATGTCGAGCGC 311

```

```

QY      303 TGTGCGCGCACTGCTGCTCTCTGGGGCTGTGTGTACTGTGTGCTTGACCTGGCCGCGC 362
Db      312 TGTGCGCGCGCGCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 371

QY      363 TGGCGCAAGACTCTGCGCGCTCTTTCAGACCAAGTTTGAAGACCGGAGCTATGACTAGC 422
Db      372 TGGCGCAAGACTCTGCGCGCTCTTTCAGACCAAGTTTGAAGACCGGAGCTATGACTAGC 431

QY      423 AGGCTGGTCTGATCTGTGGGCACTAGCCCGAGACACTGACCCCGAGCTGCGCCCTG 482
Db      432 AGGCGGGCGCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 487

QY      483 GGGCGC 487
Db      489 CCCCC 493

RESULT 4
LOCUS   B1849231
DEFINITION B1849231 541 bp mRNA linear EST 04-OCT-2001
ACCESSION B1849231
VERSION   B1849231.1 GI:15961750
KEYWORDS EST.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovidae; Bovinae; Bos.
REFERENCE Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
AUTHORS   Casas,E., Wray,J.E., White,J., Cho,J., Fahnenkrug,S.C., Bennett,
          G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
          Pette,G., Holt,I., Karaycheva,S., Liang,F., Quackenbush,J. and
          Keefe,J.W.
          Sequence evaluation of four pooled-tissue normalized bovine cDNA
          libraries and construction of a gene index for cattle
          Genome Res. 11 (4), 626-630 (2001)
JOURNAL   11282978
MEDLINE   21180013
COMMENT   Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@meat.ars.usda.gov
          Single pass sequencing. Bases called and alt trimmed with phred
          v0.980904.e. Vector identified by cross_match with the -minscore 18
          and -mismatch 12 options.
          PCR primers
          FORWARD: AGGAACAGCTATGACCAT
          BACKWARD: GTTTCCTGATGACGACG
          Plate: 100 row: H column: 8
          Seq primer: ATTAGGTGACACTATAG.
FEATURES
SOURCE    Location/Qualifiers
          1..541
            /organism="Bos taurus"
            /mol_type="mRNA"
            /db_xref="taxon:9913"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /clone_id="MARC 2B0V"
            /note="Vector: PCMV SPOR6; Site 1: NotI; Site 2: SalI;
          library made from pooled tissue from testis, thymus,
          semitendinosus muscle, longissimus muscle, pancreas,
          adrenal, and endometrium."
BASE COUNT      87 a      214 c      151 g      89 t
ORIGIN
Query Match      51.4%; Score 334.6; DB 12; Length 541;
Best Local Similarity 84.6%; Pred. No. 3e-59;
Matches 402; Conservative 0; Mismatches 64; Indels 9; Gaps 2;

```

FEATURES	source
LOCUS	BM289220
DEFINITION	530774 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION	BM289220
VERSION	BM289220.1 GI:17998246
KEYWORDS	EST.
ORGANISM	Bos taurus (cow)
SOURCE	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.
AUTHORS	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Caeas,E., Wray,J.E., White,J., Cho,J., Fahrnenkrug,S.C., Bennett,G.L., Heaton,M.P., laegreid,W.M., Rohrer,G.A., Chitto-McKown,C.G., Pettes,G., Holt,I., Karamyheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL	Genome Res. 11 (4), 626-630 (2001)
MEDLINE	21180013
PUBMED	11282978
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTCACAGTCAGACG Plate: 144 row: M column: 3 Seq primer: ATTGATGACACTATG. Location/Qualifiers 1..434

```

/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_id="MARC 3BOV"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
library made from pooled tissue from marrow, alveolar
macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."

```

Query Match	50.6%;	Score 329.6;	DB 12;	Length 434;
Best Local Similarity	88.2%;	Pred. No. 3.1e-58;		
Matches 372;	Conservative 0;	Mismatches 44;	Indels 6;	Gaps 1

Oy		20	CCCATGGACAGGGGGTATCCCGGGAAGACACCCGGGCCCCCAAGCCAGCAAGGCG--	77
Dd		1	CCCATGGACAGTGTATCCCGGGAAGACACCCGGGCCCCCAAGGCGCGAC	60
Oy		78	---GTGCCACACAGCCCTTCACTGCGGGGCCCGGCACCCCCGGCTTGAGACCACTTG	133
Dd		61	GGCAACGCCACACAGGCCCTTACCGCTGGGGGTCTCCGCAACCCCACTTGAGACCACTTG	120
Oy		134	ATCTGTCGGGTTCAGACACCCCTTACTCTGAATCTGTTCCTTCGCGCTTCGCGCGCTG	193
Dd		121	ATCTGTCGGGTTCAGACACCCCTTACTCTGAACCTGTCTCTTCGCGCTTCGCGCGCTG	180
Oy		194	GCCCTACTCCATCAAGGCCCGGAGATCAGAAAGGTGTGTGTGA CTTGAAGCGGCCCGCGT	253
Dd		181	GCCTACTCCATCAAGGCCCGGAGACCAAGAGGTGTGTGAAGATCTGAAGCGAGCCGCGT	240
Oy		254	TTTGGCTCCAAAGCCAGAGTGTATACAATCTCTGGCCCGGAGTGGAGCGCTGTGCGCGCA	313
Oy		241	CTCGCTCCAAAGCCAGAGTGTATACAATCTGTGCAAGATGTGGCGCTGTGTCGCGCG	300
Dd		301	GTCTGTCTCTCGT	360
Oy		374	TTGTGCGGCTTCTTCAAGACCAAGTTGTATGACGGGACTATGACTGACAGGCTGGTTC	433
Dd		361	TCTGCGGCTTCTTCAAGACCAAGTTGTATGACCGGACTATGACTGACAGGCTGGTTC	420
Oy		434	TG 435	
Dd		421	TG 422	
RESULT 6				
BO207465/c				
LOCUS				
DEFINITION				
ACCESION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
COMMENT				

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9365
Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to verify it as a clone from the normalized cartilage library cDNA library preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes

FEATURES SOURCE

Location/Qualifiers
1. 720
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DY1-cno-e-14-0-UI"
/tissue_type="Cartilage"
/dev_stage="ADULT"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-DY1"
/note="Organ: Femur and Tibia; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; UI-R-DY1 is a normalized cDNA library containing the following tissue(s): Rat Cartilage from Femur and Tibia. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTATGAGACG. The Rat Cartilage tissue was provided by Dr Jeff Stevens at the University of Iowa.
TAG_LIB=UI-R-DY1
TAG_TISSUE=cartilage
TAG_SEQ=CTATGAGACG"

BASE COUNT 153 a 173 c 226 g 163 t 5 others
Query Match 45.4%; Score 295.4; DB 13; Length 720;
Best Local Similarity 79.8%; Pred. No. 4e-51;
Matches 360; Conservative 0; Mismatches 85; Indels 6; Gaps 1;

QY 1 GGAAGAGACGGCGCTGGAACCCATGACACGCGCTATCCCGCGAGAGACCCGGGCCCC 60
DB 720 GGAAGAGACGGCGCTGGAACCCATGACACCTTATCNCCTGAGAGACCCGGGCCCC 661
QY 61 CAGGCCACGACGACGGCGCT-----GCCACACAGCCCTCACTAGCGCGCGCCGACACC 114
DB 660 ATCATCCCGAAGGNTATGTCAGGCCACACGCGCCCTCTGTGTGAGACACCCGCGCCC 601
QY 115 CCGGCTCGAGACCACTTGATCTGTGCGGTGTGAGACACCTCTACTGATCTGTGTG 174
DB 600 TAGGCCACGACCACTATGCTCTGTGTGTCTTGAAGACGCTGTACTGTAATCTGTGCTG 541
QY 175 CTTGCGCTTCTGTGCGCTGACCTATCATCAAGGCCCGAGATCAAGAGGTGTGTGTA 224
DB 540 CTTGCTTTCTGTGCGCTGTGCTCACTCTGTCAAGCCCGAGACAAAGGTAAGCGCGGGA 481
QY 235 CTTGGAAGCGCGCGCGGCTTTGGCTCCAAAGCCAGTGTCTACCAACTCTCTGCGCGGAT 294
DB 480 CTTGGAAGCTGCAAAAGAGATGTGCTCCAAAGCCAAAGTGTCTACCAACTCTCTGCGTGAAT 421
QY 295 GTGAGACCTGGTGGCGGCACTGTGCTCTGTGCGGCTGTGGTGTGCTGTGCGCGCGACCT 354

DB 420 GTGAGACCTGTGTCGCCATGCTGCTCTCTGGAGCTGTGTGATCTGGCGCTTGACCT 351
QY 355 GGGCCGCTGTGCGCAAGAGACTGTGCGCGCTTCTTACAGACCAAGTTGATGACCGGACTA 414
DB 360 GTCCCGGCTAGCCAAAGACTGTGCGGCTTCTTACACCAAGTTGATGAGAGACTA 301
QY 415 TACTGACAGGCTGTGCTGTCTGATCTGGGGA 445
DB 300 TAACTGAGAGTTCTGAGCCTGTCTGACCA 270

RESULT 7

AK003363
LOCUS
DEFINITION
AK003363 739 bp mRNA linear HTC 05-DEC-2002
Mus musculus 18-day embryo whole body cDNA, RIKEN full-length
enriched library, clone:111000306 product:haemopoiesis related
membrane protein 1, full insert sequence.
AK003363
VERSION
AK003363.1 GI:12833982
KEYWORDS
HTC, CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
MEDLINE
99279253
PUBMED
10349636
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
MEDLINE
20499374
PUBMED
11042159
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Komno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamanoto, R., Watsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multichannel sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Komno, S., Yamada, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gless, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Kuechenbush, J., Schriml, L.M., Staudt, F., Suzuki, R., Tomita, M.,
Wagner, L., Watanabe, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Botfield, D., Boyle, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bul, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Mazzoni, L., Mashima, J., Mazarrelli, J., Mombert, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.,
and Hayashizaki, Y.
TITLE
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

```

MEDLINE
PUBMED
12108660
11217851
AUTHORS
5
TITLE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
REFERENCE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
AUTHORS
6 (bases 1 to 739)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
Arikawa,T., Bono,H., Caminici,P., Fukuda,S., Fukunishi,Y.,
Furuno,M., Haneagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Iwama,M.,
Kasukawa,T., Kato,H., Kawai,Y., Kojima,Y., Komoh,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
TITLE
Submitted (10-JUN-2000) Yoshinobu Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-resseqc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGACGCGCCGCACTGACATGATTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adaptor of sequence[5'
GAGAGAGACGAGATCGACAGATCAATTAATTAATTAACCCCCCCC 3']. cDNA was
cleaved with XhoI and SacI. Cloning sites, 5' end: SacI; 3' end:
XhoI. Host: SOLR.
FEATURES
Source
Location/Qualifiers
1..739
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM DB:111000306"
/db_xref="MGI:1906873"
/db_xref="taxon:10090"
/clone="111000306"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18-day embryo"
62..466
/note="haemopoiesis related membrane protein 1
(MGI:MGI:1934923, GB/AJ009781, evidence: BLASTN, 99%,
match=469)
putative"
/db_xref="MGI:1934923"
721..726
/note="putative"
739
/note="putative"
BASE COUNT
160 a 228 c 186 g 165 t
ORIGIN
Query Match 43.9%; Score 286; DB 11; Length 739;
Best Local Similarity 68.1%; Pred. No. 3.5e-49;
Matches 436; Conservative 0; Mismatches 185; Indels 19; Gaps 2
1 GGAAGAGACGCGCTGGAACCATGACACGCGCTGATCCGCGGAGAGACACCGGCGCC 60

```

Db	Accession	Version	Organism	Reference Authors Title Journal Comment
Db	40		GGAAGACACGGCGGTGAACCCATGACACTTCAATTCGCCGTGAGACCCCGGGCTCC	
QY	61		CAGGCCAGCAGAGCCGGT-----GCCACACAGCCCTCACTGCGCGCCCGCACCC	
Db	100		ATCATCCCCGAGAGCTATGCTGACGCCACACAGCCCTCTCCATGGGAAACACTCGAGCC	
QY	115		CCGCGCCCGAGACCACTTGAATCTGTGTGGAGTTTACAGACCCCTCAACCTGAATCTGTGTG	
Db	160		TACACCAAGATACATGCTCTGTGTCTGTCTTACAGACATGTACCTGAATCTGTGTG	
QY	175		CTTGCGCTTCCTGCGCTGGCTTACTCATCAAGGCCCGAGATCAAGAGGTGTGTGTGA	
Db	220		CTTTGATCTCTGGCGCTGGCTGCACTGTCTCAAGGCCCGAGACTGAAGATGGCTGGGAA	
QY	235		CTGGAAGCGGCGCGGGGTGGTGGCTCCAAAGCCAGGTCAACAATCTGGCGCGAT	
Db	280		CTTGGAAGGCTGCAAGGAGTATGCTCCAAAGCCAGGTCAACAATCTGGCTGCAAT	
QY	295		GTGAGACCTGATGCGCGCTGCTGCTCTGGGCTGGTGTGACTGTGCTTCGACCT	
Db	340		GTGGAATTTGGTGTGCCCCATTGCTGCTCTGGGACTGGTGTGACTGGCGCTTGACCT	
QY	355		GGCGCGGCTGGCCAGGACTCTGCGCGCTTTTACGACCAAGTTTGAATGAGCGGACTA	
Db	400		GTCCAAATTTGACCAAAAGCTCTGGGCTTTCTTAGACAACAAAGTTTGAATGAGGACTA	
QY	415		TGACTGACAGGCTGGGCTGATCTGGGAGCACTAGCCCAAGACACTAACCCAGGCTGC	
Db	460		TAACTAAGATTCGAGACTG-----TCCCTGAACGAGGACCAACATGTCA	
QY	475		TGCCCTGGGAGCCCAATATGACTGCCCGGAGCCTGGCGCTCTTGTGTGGGCTTCAT	
Db	507		GTGATGCTGTGGCCAAACACAGCTCTGGGGTTACAGCTATATAGCACTCATATCC	
QY	535		CCCGGCCCATTCGTATCTGGGGCCCTCCAGCCCAACATGGGCACTTAAGCTGAAC	
Db	567		TGCCAGCCGTGACTGACTGAGAGCGGGCTCTTGTCCAGATGTATTGATGCTGCC	
QY	595		AGTCAGACCCCGGGGTCTTCAACCTTAACCCGAGAATTCC	
Db	627		ACTCAGATCCCTCAGACTTAACCTTAACCTAGAGGCTC	
RESULT 8				
CHS90125				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				

Oy 175 CTTGGGCTTCTGGGCTGGGCTTCTCATGAGGCGGAGATGAGAGGTGGTGTGA 234
 Db 231 CTTGGATTTCTGGGGCTGGGCTGCTGCTGACAGGCGGAGCAGAGAGATGGCTGGAA 290
 Oy 235 CTTGAGAGCGGCGGCGGCTTTTGGTCCAAAGCCAAAGTCTTCAACATCTGGCGGCAT 294
 Db 291 CTTGAGAGCTTCAAGGCAAGTGTGCTCCAAAGCCAAAGTCTTCAACATCTGGCGGCAT 350
 Oy 295 GTGAGCGTGTGGTGGCGGCTGCTGCTGAGGCGGCTGTGGTGTGACTGGTGGCTGACCT 354
 Db 351 GTGAGCATTTGGTGGCGGCTGCTGCTGCTGAGGCTGTGGTGTGACTGGCGCTTGACCT 410
 Oy 355 GGGCGGCTGGCCAAAGACTCTGCGGCTTTCTTTCAGCAACAAATTGATGAAGCGGACTA 414
 Db 411 GTCCAAAGTTAGCCAAAGACTCTGGCGGCTTTCTTTCAGCAACAAATTGATGAAGGAGACTA 470
 Oy 415 TGACTGACAG 424
 Db 471 TTAAGTAAGAG 480

RESULT 10

BY703620

LOCUS BY703620 RIKEN full-length enriched, 18-day embryo whole body Mus

DEFINITION BY703620 CDNA clone 1110003J06 5', mRNA sequence.

ACCESSION BY703620

VERSION BY703620.1 GI:27114731

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 683)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Okazaki, Y., Purano, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nishida, I., Oshino, N., Saito, R., Suzuki, H., Yamana, K., Kiyosawa, H.,
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
 Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A.,
 Quackenbush, J., Schriml, L. M., Kanapin, A., Matveeva, H., Batalov, S.,
 Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chochia, H., Corbett,
 L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,
 A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,
 Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M.,
 King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
 P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perce, G.,
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,
 Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
 B. Z., Ringwald, M., Sander, A., Schneider, C., Sempile, C. A., Setou,
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
 R. D., Tomita, M., Verardo, R., Wagner, L., Wahlefeldt, C., Wang, Y.,
 Watanabe, Y., Wells, C., Wilming, L. G., Wyshaw-Boris, A., Yangisawa,
 M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A.,
 Carinini, P., Hayatsu, N., Hirozane-Kishikawa, T., Komori, H., Nakamura,
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
 Y., Itoh, M., Kagawa, I., Yasunishi, A., Yoshino, M., Watanabe, R., Lander,
 K., Shingawa, A., Yasunishi, A., Yoshino, M., Watanabe, R., Lander,
 E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 22354683

COMMENT 1246851

Contact: Yoshitake Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saitama-cho, Tsukuba-shi, Ibaraki, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

FEATURES
 SOURCE location/qualifiers
 1..683
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1110003J06"
 /issue_type="whole body"
 /dev_stage="18-day embryo"
 /clone_id="RIKEN full-length enriched, 18-day embryo whole body"
 further details.

BASE COUNT 150 a 205 c 174 g 154 t
 ORIGIN
 Query Match 43.5%; Score 283; DB 14; Length 683;
 Best Local Similarity 79.4%; Pred. No. 1.4e-48;
 Matches 350; Conservative 0; Mismatches 85; Indels 6; Gaps 1;

Oy 1 GGAAGAGCGGCGGTGAGACCAATGACAGCGGTATCCCGGAGAGACCGGGCCCC 60
 Db 40 GGAAGAGCGGCGGTGAGACCAATGACAGCGGTATCCCGGAGAGACCGGGCCCC 99
 Oy 61 CACGCCAGAGAGCGCGGT-----GCCACACAGCGGCTTCACTGAGCGGCGGACCC 114
 Db 100 ATATCCCGAGAGCGGTATGATGAGCGGCGGACCGGCTTCCATGAGGAGACCGTGGCC 159
 Oy 115 CCGGCTGAGAGCACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 174
 Db 160 TACACACAGAGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 219
 Oy 175 CCGGCGTTCCTGGCGGTGAGCTTACATCAAGAGCGGCGGAGATGAGAGGTTGGTGA 234
 Db 220 CTTGATTTCTGGCGGTGAGCTTACATCAAGAGCGGCGGAGATGAGAGGTTGGTGA 279
 Oy 235 CTTGAGAGCGGCGGCGGCTTTTGGCTCAAGAGCGGAGTGTACAACTCTGGCGGAT 294
 Db 280 CTTGAGAGCGGCGGCGGAGTGTGTCTCAAGAGCGGAGTGTACAACTCTGGCGGAT 339
 Oy 295 GTGAGCGTGTGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
 Db 340 GTGAGCGTGTGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 399
 Oy 355 GGGCGGCTGGCCAAAGACTCTGCGGCTTTCTTTCAGCAACAAATTGATGAAGCGGACTA 414
 Db 400 GTCCAAAGTTAGCCAAAGACTCTGGCGGCTTTCTTTCAGCAACAAATTGATGAAGGAGACTA 459

QY 415 TGACTGACAGGCTGGCTCTG 435
Db 460 TAACTAGAGTTCGAGCCTG 480
RESULT 11
LOCUS B0192887/c 725 bp mRNA linear EST 30-APR-2002
DEFINITION UI-R-DRI-clk-m-07-0-UI-s1 UI-R-DRI Rattus norvegicus cDNA clone
ACCESSION B0192887
VERSION B0192887.1 GI:20368438
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE Bunkayova; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
JOURNAL Rattus.
MEDLINE 1 (bases 1 to 725)
PUBMED Bunkayova, M.F., Lennon, G. and Soares, M.B.
COMMENT Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 normalized osteoblast library cDNA library Preparation: M.B. Soares
 Lab Clone distribution: clones will be available through Research
 Genetec (www.reagen.com)
 Seq primer: M13 Forward
 POLYA=Yes.
FEATURES
 Location/Qualifiers
 1..725
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-DRI-clk-m-07-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-DRI"
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-DRI
 library is a normalized Rat Osteoblast library (nRO)
 constructed in pTZ17 vector according to the procedure
 described by Bunkayova, Lennon & Soares (Normalization and
 Subtraction: Two Approaches to Facilitate Gene Discovery.
 Genome Research 6: 791-806, 1996). The oligonucleotide
 used to prime first strand synthesis contained the
 sequence tag AGATATCA between the Not I cloning site and
 dT18 stretch. The Rat Osteoblast tissue was provided by
 Lien & Stein of the University of Massachusetts Medical
 School.
 TAG LIB=UI-R-DRI
 TAG TISSUE=osteoblast
 TAG_SEQ=AGATATCA"
BASE COUNT 155 a 175 c 228 g 165 t 2 others
ORIGIN
 Query Match 43.0%; Score 280; DB 13; Length 725;
 Best Local Similarity 79.4%; Pred. No. 6.1e-48;

Matches 358; Conservative 0; Mismatches 86; Indels 7; Gaps 2;
QY 1 GGAGAGACGGCGCTGGAACCCATGACACGGCGATATCCCGGAGACACCGGGCCCC 60
Db 719 GGAGAGAGGGGGCGCTGGAACCCATGACACCTTATATCCCGGAGAGACACCGGGCCCC 660
QY 61 CACGCCACGACGAGCGCGT-----GCCACACAGCCCTTCACACTGGCGGCCGCCACCC 114
Db 659 ATCATCCCGCCAGAGCTGATGATGACACCCACAGCGCCCTCTGTGTG-GACACCGGCC 601
QY 115 CCGCGCTCGAGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 174
Db 600 TACGCCACGACACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
QY 175 CCGCGCTCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 234
Db 540 CCGTGTCTTCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 481
QY 235 CCGTGTGTGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 294
Db 480 CCGTGTGTGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 421
QY 295 GTGAGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 354
Db 420 GTGAGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 361
QY 355 GGCCCGCTGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 414
Db 360 GTCCCGCGCTGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 301
QY 415 TGACTGACAGGCTGGCTCTG 445
Db 300 TAACTAGAGTTCGAGCCTG 270
RESULT 12
LOCUS CB321093 738 bp mRNA linear EST 04-MAR-2003
DEFINITION AGENCOUNT 12238330 NIH MGC 136 Mus musculus cDNA clone
ACCESSION CB321093
VERSION CB321093.1 GI:28845328
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Bunkayova; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL 1 (bases 1 to 738)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. David Rowe
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM327 row: a column: 01
 High quality sequence stop: 592.
FEATURES
 Location/Qualifiers
 1..738
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30288864"
 /tissue_type="embryonic limb, maxilla and mandible"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 136"
 /note="Vector: pCMV-SPORT6.1.cdb; Site 1: EcoRV; Site 2:
 NotI; Normalized, full-length enriched library from pool

QY	495	GACATCCCGAGAGCGCTGGGCGCTCTTCGTGGGGGCTCCATCCCGCATTCTGATCT	554
Db	468	AGCTCTCTGGGGATTACAGCTCTACTATACACTCTGCATCTCTGGCCAGCGTACTCAGA	527
QY	555	GGGGCGCTTCAGCCCCAATGAGGACCGTAAGGTGAAACCATGAGACCCGGGGTCTTC	614
Db	528	GGCGGCGCTCTTGTTCACGATGTGTATTGATGTGCGCCCACTAGATCTCTCAGACTTA	587
QY	615	ACCGTAAACCGAGAGTTCCC	634
Db	588	ACTTATACCTAGAGGGTCC	607

LOCUS	DEFINITION	AA000766	RESULT 14
AA000766	443 bp mRNA linear EST 18-JUL-1996		
cg38a01.r1	Soares mouse embryo NDME13.5 14.5 Mus musculus cDNA		
IMAGE:426024 5'	similar to SW:INIB.P26376		
INTERFERON-INDUCIBLE PROTEIN. [1] ;	mRNA sequence.		
AA000766			

ACCESSION	AA000766
VERSION	AA000766.1
KEYWORDS	GI:1436846
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

Euariyote; Metazoa; Chordata; Craniata; Vertebrata; Suteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 443)

REFERENCE AUTHORS
Matta, M., Hillier, L., Allen, M., Bowles, M., Dieckrich, N., Dubuque, T.,

TITLE
JOURNAL
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wyle, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The Washu-HMMI Mouse EST Project
Unpublished

COMMENT
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watscn.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:260576

```

FEATURES
  Seq primer: ETPrimers
  High quality sequence stop: 354.
  Location/Qualifiers
    1..443
    source

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:426024"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="MDH10B"
/clone_lib="Scares mouse embryo NBM13.5 14.5"
/note="Vector: pYT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer 15',
TGTATACCACTCTGAAGTGGAGCGCGCGCGGAAATTTTTTTTTTTTTTTT
T 3', on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
State Univ., from 2 ] double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pYT73 vector. Library was through one round of
normalization, and was constructed by Bento Soares and
M.Fatima Bonaldo."

```

Query Match	40.2%	Score 261.6	DB 9	Length 443
Best Local Similarity	78.5%	Pred. No. 3.4e-44		
Matches 328; Conservative	0	Mismatches 84	Indels 6	Gaps 1

QY	24	TGAGCAACGCGGTATATCCCGCGAAGACACCCGCGGCCCCACGCGCCAGCAAGGCCGGT----	79
Dp	3	TGAGCACTTATATATCCCGGTGATGATACCCCGGGCTCCATCATATCCGCAAGGCTGATGCTG	62
QY	80	--GCCACACAGCCCTTCAACTGGCGGCCCCCGACCCCCCGGCTCGAGACCACTTGATCT	137
Dp	63	CAGCCACACAGGCCCTTCCATGGGAACACTGGCCCTTACACACAGAGATCAATGCTCT	122
QY	138	GGTCGGGTTCACAGACCCCTTACTGATCTGTGTGGCTTCGCGCTTCTCGGCGCTGCGCT	197
Dp	123	GGTGTGCTTCACAGACGATGTACTGATCTGTGTGCTGCTCTTGATTCCTGGCGCTGCTCC	182
QY	198	ACTTCATCAAGGCCCGGAGATCAAGAGTGTGTGATCTGGAAAGCGGCCCGGCGCTTTTG	257
Dp	183	ACTGTGCAAGGCCCGGAGACAGAAAGTGTGCTGGAACTTGGAAAGCTGCACAAAGCAGTATG	242
QY	258	GCTCCAAAGCAGGTGTACAACATCCTGGCCGGAGTGGACGCTGTTGGCGCACATGC	317
Dp	243	GCTCCAAAGCAGGTGTACAACATCCTGGCTGCAATGTGTGACATTTGGTGGCCCCCATTTGC	302
QY	318	TGCTCCCTGGGGCTGATGTGACTGATGTCGCCGTGCACCTGGCCCGGCTGGCCCAAGGACTCTG	377
Dp	303	TGCTCTGGGACTGTGTGTACTGTGGCCCTTGCACCTGTCCAAGTTAACCACAAAGACTCTG	362
QY	378	CCGCTTCCTTCACACACCAAGTTGATGATCCGAGCTATGACTGACAGGCTGGGCTCTG	435
Dp	363	CGGCTTCTTCACACCACAATTTGATGAGAGAGACTATTACTAAGATGTCCGAGGCTG	420

RESULT 15	
W75342	
LOCUS	
DEFINITION	462 bp mRNA linear EST 20-JUN-1996
	me49n03.x1 Soares mouse embryo NbME3.5 14.5 Mus musculus cDNA
	IMAGE:390869 5' similar to SW:INIB RAT P26376
	INTERFERON-INDUCIBLE PROTEIN. [1] ", mRNA sequence.
ACCESSION	W75342.1 GI:1385557
VERSION	EST.
KEYWORDS	
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

REFERENCE	Bukaryoca; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 462) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theiling,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE	The Washu-HM1 Mouse EST Project
JOURNAL	Unpublished
COMMENT	Contact: Marra M/Mouse EST Project Washu-HM1 Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 266 1800 Fax: 314 266 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LINT; contact the IMaGRC Consortium (info@image.lnl.gov) for further information. MGI:242701 Seq primer: mob.REGA+RT High quality sequence stocp: 343. location/Qualifiers 1..462 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090"
FEATURES	
source	

```
/clone="IMAGE:390869"  
/sex="unknown"  
/tissue_type="embryo"  
/dev stage="13.5-14.5dpc total fetus"  
/lab host="DH10B"  
/clone.lib="Soares mouse embryo NMME13.5.14.5"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo (dT) primer [5',  
TGTTCACCAATCTGAAGTGGAGCGCGCCGGAATTTTATTTTATTTT  
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2  
14.5dpc embryos (total RNA provided by Minoru Ko, Wayne  
State Univ., from 2); double-stranded cDNA was ligated to  
Eco RI adaptors (Pharmacia), digested with Not I and  
cloned into the Not I and Eco RI sites of the modified  
pT73 vector. Library went through one round of  
normalization, and was constructed by Bento Soares and  
M. Fatima Bonaldo."
```

```
BASE COUNT      101 a      141 c      123 g      97 t  
ORIGIN
```

```
Query Match      40.2%; Score 261.6; DB 14; Length 462;  
Best Local Similarity 79.2%; Pred. No. 3.4e-44;  
Matches 350; Conservative 0; Mismatches 84; Indels 8; Gaps 3;
```

```
QY      1 GGAAGACGCGGCTGTGAACCCATGAGACAGCGGTATCCCGCGAGACACCCGGGCCCC 60  
DB      5 GGAAGACACGCGGCTGTGAACCCATGAGACATTCATATCCCGTGA-GACCCCGGGCTCC 63  
QY      61 CACGCGCAGCAAGCGCGT-----GCCACAGAGCCCTCAACCTGGCGGCGCCGCAACC 114  
DB      64 ATCATCCCGAAGGTGATGCTGACAGCCACAGCCCTCTCATAGGAAACCTGGGCC 123  
QY      115 CCCGCTCGAAGACACTGATCTGTGCGATGTTCAAGACCCCTACTGAATCTGTGTTG 174  
DB      124 TACACCAAGATGATCATGCTGTGCTGTCTTCAGACAGATGATCTGAATCTGTGCTG 183  
QY      175 CCTGGCTTCTGTGGCTGTGCTCACTCAATCAAGGCCGAGATCGAAGAGTGTGTGA 234  
DB      184 CTTGGAATTCCTGGCGCTGTCTCACTGTCAAGGCCGAGACAGAAATGCTGGGAA 243  
QY      235 CCTGGAAGCGCGCGCGCTTTGGGCTCCAAA-GCGAAGTCTACAAACATCTGGCGCGGA 293  
DB      244 CTTGGAAGCTGAAAGCGATGATGCTCCAAAGGCCAAGTGTACAAACATCTGGCTGCA 303  
QY      294 TGTGACGCTGTGTGCGCCACTGTCTCTGTGGGCTGTGTGACTGTGCTCTGCACC 353  
DB      304 TGTGACATTTGGTGTCCCATTTGCTGTCTGTGGACTGTGTGTGACTGGCGCTTGCAAC 363  
QY      354 TGGCCCGGCTGGCCAGAGACTGTGCGGCTTTCTTCAAGACCAAGTTGATGACGGGACT 413  
DB      364 TGTCCAACTTAGCCAAAGACTCTGCGGCTTTCTTCAAGACCAAGTTGATGAGGAGACT 423  
QY      414 ATGACTGACAGGCTGGTCTTG 435  
DB      424 ATTAATAAGATTCCGAGCTTG 445
```

```
Search completed: January 31, 2004, 15:13:27  
Job time : 2556 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2004, 13:38:24 ; Search time 71 Seconds
(without alignments)
4047.048 Million cell updates/sec

Title: US-09-914-815A-10

Perfect score: 651
Sequence: 1 ggaagagagcgcgcctggac.....cccgagcctaactctgcc 651

Scoring table:
IDENTITY_NUC
Gapop 10.0, Gapept 1.0

Searched: 56978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfillseq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93.2	14.3	642	4 US-09-370-838-119	Sequence 119, App
2	69.6	10.7	352	4 US-09-702-705-1204	Sequence 1204, App
3	69.6	10.7	352	4 US-09-736-457-1204	Sequence 1204, App
4	68	10.4	207	4 US-09-702-705-998	Sequence 998, App
5	68	10.4	207	4 US-09-736-457-998	Sequence 998, App
6	64.6	9.9	907	4 US-09-620-312D-652	Sequence 652, App
7	45.6	7.0	1311	4 US-09-252-991A-14546	Sequence 14546, A
8	45.6	7.0	3396	4 US-09-252-991A-14676	Sequence 14676, A
9	45.6	7.0	3444	4 US-09-252-991A-15078	Sequence 15078, A
10	45.2	6.9	510	4 US-09-252-991A-14818	Sequence 14818, A
11	43.2	6.6	745	1 US-08-036-555B-163	Sequence 163, App
12	43.2	6.6	745	1 US-08-469-569-163	Sequence 163, App
13	43.2	6.6	745	1 US-08-249-322A-163	Sequence 163, App
14	43.2	6.6	745	1 US-08-469-526A-163	Sequence 163, App
15	43.2	6.6	745	2 US-08-734-651A-163	Sequence 163, App
16	43.2	6.6	745	2 US-08-469-660-163	Sequence 163, App
17	43.2	6.6	745	3 US-08-341-018-51	Sequence 51, App1
18	43.2	6.6	745	3 US-08-470-335-163	Sequence 163, App
19	43.2	6.6	745	3 US-08-735-021-163	Sequence 163, App
20	43.2	6.6	745	3 US-08-734-664A-163	Sequence 163, App
21	43.2	6.6	745	3 US-08-470-339-163	Sequence 163, App
22	43.2	6.6	745	4 US-08-467-602-163	Sequence 163, App
23	43.2	6.6	745	5 PCT-US94-05083C-159	Sequence 159, App
24	43.2	6.6	745	5 PCT-US95-06846A-163	Sequence 163, App
25	43.2	6.6	2003	1 US-08-036-555B-21	Sequence 21, App1
26	43.2	6.6	2003	1 US-08-469-569-21	Sequence 21, App1
27	43.2	6.6	2003	1 US-08-249-322A-21	Sequence 21, App1

28	43.2	6.6	2003	1 US-08-469-526A-21	Sequence 21, App1
29	43.2	6.6	2003	2 US-08-734-591A-21	Sequence 21, App1
30	43.2	6.6	2003	2 US-08-469-660-21	Sequence 21, App1
31	43.2	6.6	2003	3 US-08-341-018-71	Sequence 71, App1
32	43.2	6.6	2003	3 US-08-470-335-21	Sequence 21, App1
33	43.2	6.6	2003	3 US-08-735-021-21	Sequence 21, App1
34	43.2	6.6	2003	3 US-08-734-664A-21	Sequence 21, App1
35	43.2	6.6	2003	3 US-08-470-339-21	Sequence 21, App1
36	43.2	6.6	2003	4 US-08-467-602-21	Sequence 21, App1
37	43.2	6.6	2003	5 PCT-US94-05083C-21	Sequence 21, App1
38	43.2	6.6	2003	5 PCT-US95-06846A-21	Sequence 21, App1
39	41.2	6.3	764	4 US-08-991-789A-283	Sequence 283, App
40	41.2	6.3	764	4 US-09-062-451-283	Sequence 283, App
41	41.2	6.3	764	4 US-09-289-198-283	Sequence 283, App
42	40.6	6.2	2846	3 US-09-613-182-5	Sequence 5, App1
43	39.8	6.1	2616	4 US-09-252-991A-16042	Sequence 16042, A
44	39.8	6.1	2739	4 US-09-252-991A-16558	Sequence 16558, A
45	39.4	6.1	615	4 US-09-501-115-21	Sequence 21, App1

ALIGNMENTS

```
RESULT 1
US-09-370-838-119
; Sequence 119, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadh
; APPLICANT: Sezarist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.47501
; CURRENT APPLICATION NUMBER: US/09/370, 838
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-370-838-119

Query Match      14.3%; Score 93.2; DB 4; Length 642;
Best Local Similarity 66.3%; Pred.No. 6.8e-13;
Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY      125 GACCACTTATCTGTGCTGCTTCAAGCAACCTTACCTGATCTGTGCTTGCCTTGCCTTC 184
      |||||
DB      216 GACCACTTATCTGTGCTTCAAGCAACCTTACCTGATCTGTGCTTGCCTTGCCTTTC 275

QY      185 CTGGCGCTGCTTCTTCAATCAAGCCGAGATCAGAAAGTGTGCTTGAATCTTGAAGCG 244
      |||||
DB      276 ATAGCACTTGTCTTCTTCAAGCTTGAAGCTTGAAGCAAGAAATGTGCTTGAAGCAAGCG 335

QY      245 GCCCGCGCTTGTGCTTCAAGCAAGCTTCAATCTGATCTGCGCGCGATGTGCAAGCGT 304
      |||||
DB      336 GCCCGCGCTTGTGCTTCAAGCAAGCTTCAATCTGATCTGCGCGCGATGTGCAAGCGT 395

QY      305 GTGCGCGCACTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 326
      |||||
DB      396 CTCATGACCATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 417

RESULT 2
US-09-702-705-1204
; Sequence 1204, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
```

```
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1204
LENGTH: 352
TYPE: DNA
ORGANISM: Homo sapien
US-09-702-705-1204
```

```
Query Match 10.7%; Score 69.6; DB 4; Length 352;
Best Local Similarity 56.6%; Pred. No. 1.6e-07;
Matches 129; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
```

```
QY 23 ATGACACGCGGTATCCCGGAGACACCCGCGCCCGACGAGAGCCGCGTCC 82
DB 123 ATGACACGAGGAACATGAGGTGCTGTGGGGGACCCCGACGACATCTTCCA 182
QY 83 CACACAGCCCTCACACTGGCGGCCCCCGGCTTCGAGACATCTATCTGTG 142
DB 183 AGGTCCACCGTATCAATCCACGAGACCTCGGCCGACATGCTGTGTC 242
QY 143 GTGTCAGACCCCTTACCTGATCTGTGCTCTGCGCTTCGCGCTGCTACTCC 202
DB 243 CTGTTCAACCCCTTCTTCTGACGCTGTCTGTGGGCTTATGACATCGCTACTCC 302
QY 203 ATCAAGCCCGAGATCAGAGGTGTGTGTAAGCTGGAAGCGCGCCG 250
DB 303 GTGAAGTCTAGGAGACAGAAAGTGTGGCGAGCTGACCGGGGCCAG 350
```

RESULT 3

```
US-09-736-457-1204
Sequence 1204, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1204
LENGTH: 352
TYPE: DNA
ORGANISM: Homo sapien
US-09-736-457-1204
```

```
Query Match 10.7%; Score 69.6; DB 4; Length 352;
Best Local Similarity 56.6%; Pred. No. 1.6e-07;
```

```
Matches 129; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 23 ATGACACGCGGTATCCCGGAGACACCCGCGCCCGACGAGAGCCGCGTCC 82
DB 123 ATGACACGAGGAACATGAGGTGCTGTGGGGGACCCCGACGACATCTTCCA 182
QY 83 CACACAGCCCTCACACTGGCGGCCCCCGGCTTCGAGACATCTATCTGTG 142
DB 183 AGGTCCACCGTATCAATCCACGAGACCTCGGCCGACATGCTGTGTC 242
QY 143 GTGTCAGACCCCTTACCTGATCTGTGCTCTGCGCTTCGCGCTGCTACTCC 202
DB 243 CTGTTCAACCCCTTCTTCTGACGCTGTCTGTGGGCTTATGACATCTGCTACTCC 302
QY 203 ATCAAGCCCGAGATCAGAGGTGTGTGTAAGCTGGAAGCGCGCCG 250
DB 303 GTGAAGTCTAGGAGACAGAAAGTGTGGCGAGCTGACCGGGGCCAG 350
```

```
RESULT 4
US-09-702-705-998
Sequence 998, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
```

```
APPLICANT: Wang, Tongtong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 998
LENGTH: 207
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(207)
OTHER INFORMATION: n = A,T,C or G
US-09-702-705-998
```

Query Match

```
Best Local Similarity 71.8%; Score 68; DB 4; Length 207;
Pred. No. 3.4e-07;
Matches 89; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 125 GACCATTTGATCTGTGTGTTGTTGACACCTCTTACTGAATCTGTGTGCTTCC 184
DB 83 GACCATGTCTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 142
QY 185 CTGCGCGGTGCTTCTCTCATCAGAGCCGAGATCAGAGGTGTGTGTAAGCTGGAAGCG 244
DB 143 ATAGCATTCGCTTCTCTCTGTAAGTCTAGGAGACAGAAAGTGTGGCGAGCTGACCGGG 202
QY 245 GCCC 248
DB 203 GCCC 206
```

RESULT 5

```
US-09-736-457-998
Sequence 998, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
```



```
RESULT 11
US-08-036-555B-163
; Sequence 163, Application US/08036555B
; Patent No. 5530109
; GENERAL INFORMATION:
; APPLICANT: Goodheart, Andrew; Stroobant, Paul;
; APPLICANT: Minshetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Malo Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; TITLE OF INVENTION: Preparation and Use
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 435
; APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-036-555B-163

Query Match 6.6%; Score 43.2; DB 1; Length 745;
Best Local Similarity 57.4%; Pred. No. 0.21;
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 246 CCCGCGCTTTGGCTCCAAAGCCAAAGTGTACAAATCTCGGCGCGATGTGACGCTGG 305
DB 29 CCGGGGCTCCCGGCGCCCGGCGCCAGCGCCCGGCTCGCGCGCTCGTCCGCGCGC 88
QY 306 TCGCGCACGTGCTGCTCTTGAGGCTGTGTGTAAGTGTGCTGCACTGAGCCCGGCTGG 365
DB 89 TCGCGCTGCTGCGCACTGCTGCTGTGCGGACCGCGGCTTGGCGCGGAGCGCGG 148
QY 366 CCAAGACTGCGCGC 381
```

```
DB 149 CCGGCAAGAGCGCGC 164

RESULT 12
US-08-469-569-163
; Sequence 163, Application US/08469569
; Patent No. 5606032
; GENERAL INFORMATION:
; APPLICANT: Goodheart, Andrew; Stroobant, Paul;
; APPLICANT: Minshetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Malo Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; TITLE OF INVENTION: Preparation and Use
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,569
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 745
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-469-569-163

Query Match 6.6%; Score 43.2; DB 1; Length 745;
Best Local Similarity 57.4%; Pred. No. 0.21;
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 246 CCCGCGCTTTGGCTCCAAAGCCAAAGTGTACAAATCTCGGCGCGATGTGACGCTGG 305
DB 29 CCGGGGCTCCCGGCGCCCGGCGCCAGCGCCCGGCTCGCGCGCTCGTCCGCGCGC 88
QY 306 TCGCGCACGTGCTGCTCTTGAGGCTGTGTGTAAGTGTGCTGCACTGAGCCCGGCTGG 365
DB 89 TCGCGCTGCTGCGCACTGCTGCTGTGCGGACCGCGGCTTGGCGCGGAGCGCGG 148
```

Qy 366 CCAAGACTCTGCCG 381
Db 149 CCGGCAACGAGCGGC 164

RESULT 13

US-08-249-322A-163
Sequence 163, Application US/08249322A
Patent No. 5716930

GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew; Stroobant, Paul;
Minghetti, Luisa; Waterfield, Michael; Marchionl, Mark;
APPLICANT: Chen, Malo Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pelte & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/249,322A

FILING DATE: 26-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/965,173

FILING DATE: 23-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/907,138

FILING DATE: 30-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/863,703

FILING DATE: 03-APRIL-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.K. 91 07566.3

FILING DATE: 10-APRIL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Tsai, Christine H.

REGISTRATION NUMBER: 34,266

TELEPHONE/DOCKET NUMBER: LUD 250.4

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 163:

SEQUENCE CHARACTERISTICS:

LENGTH: 745

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-249-322A-163

Query Match 6.6%; Score 43.2; DB 1; Length 745;

Best Local Similarity 57.4%; Pred. No. 0.21; Mismatches 58; Indels 0; Gaps 0;

Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 306 TGCCGCACTGCTCTCTGCGGCTGTGTGACTGTGCTGCACCTTGCCCGCTGG 365
Db 89 TGCCGCTGCTGCACCTGCTGCTGTGCGGACCGCGGCTCTGCGGCGCGCG 148

Qy 366 CCAAGACTCTGCCG 381
Db 149 CCGGCAACGAGCGGC 164

RESULT 14

US-08-469-526A-163
Sequence 163, Application US/08469526A
Patent No. 5792849

GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew
APPLICANT: Stroobant, Paul
APPLICANT: Minghetti, Luisa
APPLICANT: Waterfield, Michael
APPLICANT: Marchionl, Mark
APPLICANT: Chen, Malo Su
APPLICANT: Hiles, Ian
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
NUMBER OF SEQUENCES: 187
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Biring LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,526A

FILING DATE: 06 June 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

APPLICATION NUMBER: 07/965,173

FILING DATE: 23-OCT-1992

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

APPLICATION NUMBER: 07/907,138

FILING DATE: 03-JUN-1992

APPLICATION NUMBER: 07/863,703

FILING DATE: 03-APRIL-1992

APPLICATION NUMBER: U.K. 91 07566.3

FILING DATE: 10-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Bleker-Brady, Kristina

REGISTRATION NUMBER: 39,109

TELEPHONE/DOCKET NUMBER: 04585/00200A

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

INFORMATION FOR SEQ ID NO: 163:

SEQUENCE CHARACTERISTICS:

LENGTH: 745

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-469-526A-163

Query Match 6.6%; Score 43.2; DB 1; Length 745;

Best Local Similarity 57.4%; Pred. No. 0.21; Mismatches 58; Indels 0; Gaps 0;

Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 246 CCGGCGTTTGGCTCCAAAGCCCAAGTGTACAAACATCTGCGCGCGATGTGAGCGCTGG 305
Db 29 CCGGCGTCCCGGCGCCCGGCGCCAGCGCCCGGCGCTCCGCGCGCTGCTGTGCGCGCGCG 86

OY	246	CCGCGCGT	TTGGCTC	CAAAAGC	CAAGTGC	TACACAT	CTGCG	CCGCAT	TGTGAG	CGCTGG	305	
Db	29	CCGCGCGT	CCCGGCCCC	CGGCGCC	CAAGCG	CCCCGCGCT	CCGCG	CCCGCT	CTGTGTG	CGCGCGC	88	
OY	306	TGCGCGCA	CTGCTG	CTCTG	GGGCTG	GGTGGT	GATCGT	GTGC	CTTGCA	CTTGCG	CGGCGTGG	365
Db	89	TGCGCGCT	GTGCTGC	CACTACT	GTGTGT	GTCTGT	GGGAA	CCGCGG	CCCTTG	CGCTCG	GGGCGCGG	148
OY	366	CCAAGAC	CTTGCG	CGC	381							
Db	149	CCGCGCA	ACGAGCG	CGC	164							

RESULT 15-591A-163
 US-08-734-591A-163
 Sequence 163: Application US/08734591A
 Patent No. 5854220
 GENERAL INFORMATION:
 APPLICANT: Goodheart, Andrew
 APPLICANT: Stoochart, Paul
 APPLICANT: Minghetti, Luisa
 APPLICANT: Waterfield, Michael
 APPLICANT: Hiles, Ian
 APPLICANT: Marchionni, Mark
 APPLICANT: Chen, Mario
 TITLE OF INVENTION: GIAL MITOGENIC FACTORS, THEIR
 TITLE OF INVENTION: PREPARATION AND USE
 NUMBER OF SEQUENCES: 187
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Clark & Ebling LLP
 STREET: 176 Federal Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM Compatible Pentium
 OPERATING SYSTEM: Windows95
 SOFTWARE: WordPerfect (Version 7.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/734,591A
 FILING DATE: 22-OCT-1996
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/470,335
 FILING DATE: 06-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/036,555
 FILING DATE: 03-MAR-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/965,173
 FILING DATE: 23-OCT-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/940,389
 FILING DATE: 03-SEP-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/907,138
 FILING DATE: 30-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/863,703
 FILING DATE: 03-APR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: UK 91 07566.3
 FILING DATE: 10-APR-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Bleker-Brady, Kristina
 REGISTRATION NUMBER: 39,109
 REFERENCE/DOCKET NUMBER: 04585/00200P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 428-0200
 TELEFAX: (617) 428-7045
 TELEX:

```

; INFORMATION FOR SEQ ID NO: 163
;
; SEQUENCE CHARACTERISTICS:
;
;   LENGTH: 745
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;
US-08-734-591A-163

```

Query Match	6.6%	Score 43.2	DB 2	Length 745
Best Local Similarity	57.4%	Pred. No. 0.21		
Matches 78; Conservative	0;	Mismatches 58;	Indels 0;	Gaps 0;

Qy	246	CCGGGGGTTTGGCTCCAAAGCCAAAGTCTAACACATCTGCGCCGCAATGAGAGCGCTGG	305
Db	29	CCGGGGGTTTCCCGACCCCGGGCCACAGCCGCCCGGCTCCCGCGCCCGCTGTCGCGCGCGC	88
Qy	306	TGCGGCACTGCTGCTCTGAGGAGCTGAGTGATGAGGCGCTTGACACTGTGACCGGAGCTGG	365
Db	89	TGCGGCTGCTGCACTACTGCTGCTGCTGGGAGACGCGGCGCTTGCGGCCCGGAGGCGGCGG	148
Qy	366	CCAAGGACTTGTGCGCG	381
Db	149	CCGGCAACGAGGCGCG	164

Search completed: January 31, 2004, 15:14:48
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2004, 12:18:23 ; Search time 286 Seconds
(without alignments)
6144.525 Million cell updates/sec

Title: US-09-914-815A-10
Perfect score: 651

Sequence: 1 ggaagagacggcgtggaac.....ccggagccctaactctgcc 651

Scoring table: IDENTITY_NUC
Gap 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	651	100.0	651	21	AAA64687
2	424.6	65.2	651	24	AB189793
3	96.4	14.8	447	22	AA119241
4	95.4	14.7	326	22	ABA47169
5	95.4	14.7	326	22	ABA65053
6	95.4	14.7	326	22	ABA32159
7	95.4	14.7	326	22	AAK13477
8	95.4	14.7	326	22	AAK39215

C	9	95.4	14.7	326	22	AA170025	Probe #958 for ge
C	10	95.4	14.7	326	22	AA145223	Probe #13909 used
C	11	95.4	14.7	326	22	AA105732	Probe #5723 used t
C	12	95.4	14.7	326	23	AB388800	Human liver single
C	13	95.4	14.7	326	24	AB513299	Human genome-deriv
C	14	95.4	14.7	459	22	ABA42012	Human breast cell
C	15	95.4	14.7	459	22	ABA52433	Human foetal liver
C	16	95.4	14.7	459	22	ABA22426	Probe #692 for gen
C	17	95.4	14.7	459	22	AAK00702	Human brain expres
C	18	95.4	14.7	459	22	AAK6153	Human bone marrow
C	19	95.4	14.7	459	22	AA110783	Probe #716 for gen
C	20	95.4	14.7	459	22	AA112041	Probe #727 used to
C	21	95.4	14.7	459	22	AA100708	Probe #69 used to
C	22	95.4	14.7	459	23	AB525742	Human liver single
C	23	95.4	14.7	459	24	AB500738	Human genome-deriv
C	24	95.4	14.6	621	25	AA126291	Human breast cance
C	25	93.4	14.3	621	25	ACC42315	Human MAP kinase c
C	26	93.4	14.3	631	24	AB099282	Human coding sequ
C	27	93.4	14.3	905	24	ABK84587	Human cDNA differe
C	28	93.4	14.3	905	24	ABK69222	Human liver kinase
C	29	93.4	14.3	905	25	ACC42316	Human MAP kinase c
C	30	93.4	14.3	905	25	ACC42354	Human MAP kinase c
C	31	93.4	14.3	905	25	ACC42355	Human MAP kinase c
C	32	93.2	14.3	484	24	AB060102	Human colon cancer
C	33	93.2	14.3	606	25	ACC42317	Human MAP kinase c
C	34	93.2	14.3	621	25	ACC42319	Human MAP kinase c
C	35	93.2	14.3	642	20	AAZ07228	Human lung tumour
C	36	93.2	14.3	642	21	AAZ91198	Human lung tumour
C	37	93.2	14.3	642	21	AAZ31193	Human lung tumour
C	38	93.2	14.3	695	24	AA595009	Human DNA sequence
C	39	93.2	14.3	712	23	AA575071	DNA encoding novel
C	40	93.2	14.3	761	24	AB056182	Human ovarian anti
C	41	93.2	14.3	808	25	AAA46668	CDNA of I-BU gene
C	42	93.2	14.3	808	25	ACC42318	Human MAP kinase c
C	43	92.4	14.2	637	23	AA575776	DNA encoding novel
C	44	91.8	14.1	716	22	AA159665	Human polynucleoti
C	45	91.8	14.1	999	22	AA157879	Human polynucleoti

ALIGNMENTS

RESULT 1	
AAA64687	AAA64687 standard; cDNA, 651 BP.
ID	
XX	AAA64687;
AC	
XX	
DT	02-JAN-2001 (first entry)
XX	
DE	cDNA encoding a human leukocyte and blood related protein (LBAP).
XX	
KW	Human; leukocyte and blood related protein; LBAP; arteriosclerosis;
KW	cell proliferative disorder; acinic keratosis; atherosclerosis;
KW	burialis; cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
KW	myelofibrosis; paroxysmal nocturnal hemoglobinuria; cancer;
KW	adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;
KW	teratocarcinoma; autoimmune disorder; inflammatory disorder;
KW	acquired immunodeficiency syndrome; AIDS; Addison's disease;
KW	adult respiratory distress syndrome; allergy; ankylosing spondylitis;
KW	amyloidosis; anaemia; asthma; autoimmune haemolytic anaemia; infection;
KW	Werner syndrome; haemodialysis; extracorporeal circulation; trauma; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	23..421
FT	/tag= a
FT	/product= "leukocyte and blood related protein (LBAP)"
XX	MO200052161-A2.
XX	
PD	08-SEP-2000.

XX 29-FEB-2000; 2000WO-US05153.
 XX 01-MAR-1999; 99US-0122080.
 XX (INCY-) INCYTE PHARM INC.
 XX
 XX Lal P, Yue H, Hillman JL, Lu DM, Baughn MR, Tang YT, Azimzai Y,
 XX WPI; 2000-587310/55.
 XX P-PSDB; AAB08767.
 XX
 PT Leukocyte and blood associated proteins and polynucleotides encoding
 PT them, useful for diagnosis, treatment and prevention of
 PT autoimmune/inflammatory disorders and cell proliferative disorders
 PT including cancer -
 PS
 XX
 PS Claim 4; Page 70; 70pp; English.
 CC The present sequence encodes a human leukocyte and blood related
 CC protein, designated LBAP. LBAP polynucleotides and polypeptides are
 CC useful for treating or preventing a disorder associated with decreased
 CC expression or activity of LBAP including a cell proliferative disorder
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,
 CC cirrhosis, hepatitis, mixed connective tissue disease (MCTD),
 CC myelofibrosis, paroxysmal nocturnal hemoglobinuria, etc., cancers
 CC including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma,
 CC sarcoma, teratocarcinoma and in particular cancers of the adrenal
 CC gland, bladder, bone, bone marrow, brain, breast, cervix, etc., and
 CC an autoimmune/inflammatory disorder such as acquired immunodeficiency
 CC syndrome (AIDS), Addison's disease, adult respiratory distress syndrome,
 CC allergies, ankylosing spondylitis, amyloidosis, anaemia, asthma,
 CC atherosclerosis, autoimmune haemolytic anaemia, etc., Werner syndrome,
 CC complications of cancer, haemodialysis, and extracorporeal circulation,
 CC viral, bacterial, fungal, parasitic, protozoan, and helminthic
 CC infections, and trauma.
 CC
 XX
 XX Sequence 651 BP; 107 A; 243 C; 185 G; 116 T; 0 other;
 SQ
 Query Match 100.0%; Score 651; DB 21; Length 651;
 Best Local Similarity 100.0%; Pred. No. 1.2e-134;
 Matches 651; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAAGAGACGCGCGCTGGAACCCATGAGACGCGCGTATCCCGGAGACACCCGGGCCCC 60
 DB 1 GGAAGAGACGCGCGCTGGAACCCATGAGACGCGCGTATCCCGGAGACACCCGGGCCCC 60
 QY 61 CACGCCCGACGAGCGCGGTGCGCCACACAGCCCTTCACTGAGGCGCCCGGACCCCGCC 120
 DB 61 CACGCCCGACGAGCGCGGTGCGCCACACAGCCCTTCACTGAGGCGCCCGGACCCCGCC 120
 QY 121 TCGAGACCACTGATCTGCTGCGGTGTCAGACCCCTTCACTGATCTGCTGCTGCTGCTG 180
 DB 121 TCGAGACCACTGATCTGCTGCGGTGTCAGACCCCTTCACTGATCTGCTGCTGCTGCTG 180
 QY 181 CTTCCTGCGCTGCGCTTACTCATATAGAGCCCGAGATCAAGAGGTGTTGATGACTGGA 240
 DB 181 CTTCCTGCGCTGCGCTTACTCATATAGAGCCCGAGATCAAGAGGTGTTGATGACTGGA 240
 QY 241 AGCGGCGCGCGCTTTTGGCTTCAAGCCAGTGTCAACAATCTGCGCGCGCATGTGAC 300
 DB 241 AGCGGCGCGCGCTTTTGGCTTCAAGCCAGTGTCAACAATCTGCGCGCGCATGTGAC 300
 QY 301 GCTGTGCGCGCACTGCTGCTCTGCGGCGTGTGATGATGCTGCTGCTGCTGCTGCTG 360
 DB 301 GCTGTGCGCGCACTGCTGCTCTGCGGCGTGTGATGATGCTGCTGCTGCTGCTGCTG 360
 QY 361 GCTGTGCGCGCACTGCTGCTCTTTCAGACCAAGTTGATGACGCGGACTATGACTG 420
 DB 361 GCTGTGCGCGCACTGCTGCTCTTTCAGACCAAGTTGATGACGCGGACTATGACTG 420
 QY 421 AAGAGGTGCGCTGCTGATCTGAGGCGCTGAGCCAGAGACCTGACCCGAGCTGCGCC 480

DB 421 AAGAGGTGCGCTGCTGATCTGAGGCGCACTAGCCCGAGACACTGACCCGAGCTGCTGCCCC 480
 QY 481 TGGGGCCCAATATGACTCTCCCGGAGCGTGGACCTCTTCTGTGGGAGCTTCATCCCTGC 540
 DB 481 TGGGGCCCAATATGACTCTCCCGGAGCGTGGACCTCTTCTGTGGGAGCTTCATCCCTGC 540
 QY 541 CCATCTGATCTGAGGCGCTTCAAGCCCAACATGTGGACCTTAAGGCTGAACAGTCAG 600
 DB 541 CCATCTGATCTGAGGCGCTTCAAGCCCAACATGTGGACCTTAAGGCTGAACAGTCAG 600
 QY 601 ACCCGGGGCTTCACTTACCCGAGAGTCCCGGCGCTTCACTGCGCC 651
 DB 601 ACCCGGGGCTTCACTTACCCGAGAGTCCCGGCGCTTCACTGCGCC 651

RESULT 2
 ABL89793
 ID ABL89793 standard; cDNA; 451 BP.
 XX
 XX ABL89793;
 XX
 XX 24-MAY-2002 (first entry)
 XX
 XX Human polynucleotide SEQ ID NO 355.
 DE
 XX
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 XX vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
 XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 XX neurological disease; infection; human; secreted protein; gene; ss.
 OS
 XX Homo sapiens.
 XX
 XX WO200190304-A2.
 XX
 XX 29-NOV-2001.
 XX
 XX 18-MAY-2001; 2001WO-US16450.
 XX
 XX 19-MAY-2000; 2000US-205515P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 XX WPI; 2002-122018/16.
 XX
 XX P-PSDB; ABB9384.
 XX
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 XX prevention of neural, immune system, muscular, reproductive,
 XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 XX disorders -
 XX
 PS Claim 4; SEQ ID NO 355; 2081bp + Sequence Listing; English.
 XX
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 XX (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 XX medical conditions e.g. by protein or gene therapy. The genes are
 XX isolated from a range of human tissues disclosed in the specification.
 XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
 XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
 XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 XX colitis; (c) cardiovascular disorders such as myocardial ischaemia;
 XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 XX and parasitic infections.
 XX Note: The sequence data for this patent did not form part of the
 XX printed specification, but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 451 BP; 74 A; 156 C; 137 G; 83 T; 1 other;
Query Match 65.2%; Score 424.6; DB 24; Length 451;
Best Local Similarity 98.4%; Pred. No. 1.1e-84;
Matches 438; Conservative 1; Mismatches 5; Indels 1; Gaps 1;
OY 7 GACGGCGGTGAGAACCATGAGACAGCGGCTATCCCGGAGACACCCGGGCTCCAGAGCC 66
DB 6 GAGGGCGGTGAGAACCATGAGACAGCGGCTATCCCGGAGACACCCGGGCTCCAGAGCC 65
OY 67 CAGCAAGCCCGGTGAGAACCATGAGACAGCGGCTATCCCGGAGACACCCGGGCTCCAGAG 126
DB 66 CAGCAAGCCCGGTGAGAACCATGAGACAGCGGCTATCCCGGAGACACCCGGGCTCCAGAG 125
OY 127 CCACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186
DB 126 CCACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
OY 187 GACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246
DB 186 GACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 245
OY 247 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 306
DB 246 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305
OY 307 GCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
DB 306 GCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
OY 367 CAAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
DB 365 CAAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
OY 427 TGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 451
DB 425 TGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 449

RESULT 3
AAL19241/c
ID AAL19241 standard; cDNA; 447 BP.
XX AAL19241;
XX AC AAL19241;
XX DT 07-DEC-2001 (first entry)
XX DE Human breast cancer expressed polynucleotide 11698.
XX KW Human; breast cancer; cell marker; cytostatic; ss.
XX OS Homo sapiens.
XX PN WO200151628-A2.
XX PD 19-JUL-2001.
XX PF 10-JAN-2001; 2001WO-US00798.
XX PR 14-JAN-2000; 2000US-0176077.
XX PR 14-MAR-2000; 2000US-0189167.
XX PR 24-MAR-2000; 2000US-0192089.
XX PR 22-MAR-2000; 2000US-0193480.
XX PR 15-MAY-2000; 2000US-0205230.
XX PR 09-JUN-2000; 2000US-0211315.
XX PR 25-JUL-2000; 2000US-0220534.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PT Lilie J, Xu Y, Wang Y, Steinmann K;
XX DR WPI; 2001-451856/48.

XX FT New peptide useful as a marker for the diagnosis of breast cancer -
XX PS Claim 1; Page 2084; 3695pp; English.
XX CC The invention relates to human breast cancer expressed polynucleotides
XX CC (AAL07544-AAL26789) and methods of assessing whether a patient is
XX CC afflicted with breast cancer by examining the correlation between the
XX CC expression of certain markers and the cancerous state of breast cells.
XX CC The polynucleotides and encoded polypeptides are potential markers for
XX CC detecting, diagnosing, monitoring, characterizing treating and
XX CC potentially preventing breast cancer. The polynucleotides and encoded
XX CC polypeptides are also useful for isolating compounds with cytostatic
XX CC activity.
SQ Sequence 447 BP; 102 A; 101 C; 121 G; 123 T; 0 other;
Query Match 14.8%; Score 96.4; DB 22; Length 447;
Best Local Similarity 62.4%; Pred. No. 3.6e-12;
Matches 151; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
OY 114 CCCGCTCGAGACCACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 173
DB 421 CCGTGGCTCGAGACCACTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362
OY 174 GCTTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 233
DB 361 GCTTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
OY 234 ACCGGAAGGGGGCCCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 293
DB 301 ACCGGAAGGGGGCCCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242
OY 294 TGTGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 353
DB 241 TTTTGAACATCTTCATGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
OY 354 TG 355
DB 181 AG 180

RESULT 4
ABA47169/c
ID ABA47169 standard; DNA; 326 BP.
XX ABA47169;
XX AC ABA47169;
XX DT 01-FEB-2002 (first entry)
XX DE Human breast cell single exon nucleic acid probe #5864.
XX KW Human; microarray; single exon probe; gene expression; breast;
XX KW disease; cancer; ss.
XX OS Homo sapiens.
XX PN WO200157271-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00662.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234587.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PT

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
XX
PS Claim 4; SEQ ID NO 5864; 327bp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
Query Match 14.7%; Score 95.4; DB 22; Length 326;
Best Local Similarity 67.2%; Pred. No. 5.8e-12;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 126 ACCACTGATGTCGTGGTGTGTCAGACCCCTTACTGATGTCGTGGCTTCC 185
DB 209 ACCATGTTGTGTGTCCTGTTCAACACCCCTTCAATGAACCCCTGCTGGGCTTCA 150
QY 186 TGGCGTGGCTTACTCATCAAGCCCGAGATCAAGAGTGTGTGTCCTGGAAGCG 245
DB 149 TAGGTTACCTTACTCATGAAGTCTAGGAGACAGAAAGTGTGGCCACTGACCGGG 90
QY 246 CCCGCGTTTGGCTTCCAAAGCCCAAGTCTAACAATCTGCGCGGATGTGAAGCTGG 305
DB 89 CCCAGGCTATGCTCTCCACCGCCCAAGTGTCTGAACATCTGGGCTTGAATTGGGCATCC 30
QY 306 TGGCGCACTGCTGCTCTCTGG 326
DB 29 TCATGACCATCTCTGCTCATCG 9
RESULT 5
ABA65053/C
ID ABA65053 standard; DNA; 326 BP.
XX
XX ABA65053;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human foetal liver single exon nucleic acid probe #13358.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX Homo sapiens.
XX
XX MO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00669.
XX
XX PF
XX PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX
PS Claim 4; SEQ ID NO 13358; 639bp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
Query Match 14.7%; Score 95.4; DB 22; Length 326;
Best Local Similarity 67.2%; Pred. No. 5.8e-12;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 126 ACCACTGATGTCGTGGTGTGTCAGACCCCTTACTGATGTCGTGGCTTCC 185
DB 209 ACCATGTTGTGTGTCCTGTTCAACACCCCTTCAATGAACCCCTGCTGGGCTTCA 150
QY 186 TGGCGTGGCTTACTCATCAAGCCCGAGATCAAGAGTGTGTGTCCTGGAAGCG 245
DB 149 TAGGTTACCTTACTCATGAAGTCTAGGAGACAGAAAGTGTGGCCACTGACCGGG 90
QY 246 CCCGCGTTTGGCTTCCAAAGCCCAAGTCTAACAATCTGCGCGGATGTGAAGCTGG 305
DB 89 CCCAGGCTATGCTCTCCACCGCCCAAGTGTCTGAACATCTGGGCTTGAATTGGGCATCC 30
QY 306 TGGCGCACTGCTGCTCTCTGG 326
DB 29 TCATGACCATCTCTGCTCATCG 9
RESULT 6
ABA32159/C
ID ABA32159 standard; DNA; 326 BP.
XX
XX ABA32159;
XX
XX 23-JAN-2002 (first entry)
XX
XX Probe #10625 for gene expression analysis in human heart cell sample.
XX
XX Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
XX Homo sapiens.
XX
XX MO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX PF
XX PR

XX	PR	04-FEB-2000; 2000US-0190312.
XX	PR	26-MAY-2000; 2000US-0207456.
XX	PR	30-JUN-2000; 2000US-0608408.
XX	PR	03-AUG-2000; 2000US-0632366.
XX	PR	21-SEP-2000; 2000US-0234687.
XX	PR	27-SEP-2000; 2000US-0236359.
XX	PR	04-OCT-2000; 2000GB-0024263.
PA	(MOL-)	MOLECULAR DYNAMICS INC.
PI	Penn SG,	Hanzel DK, Chen W, Rank DR;
XX	WIPI;	2001-488899/53.
PT	Single exon nucleic acid probes for analyzing gene expression in human hearts -	
XX	Claim 4;	SEQ ID No 10625; 530bp; English.
CC	The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognostating diseases of the human heart, and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.	
CC	Note:	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ	Sequence	326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
Query Match	14.7%;	Score 95.4; DB 22; Length 326;
Best Local Similarity	67.2%;	Pred. No. 5.8e-12;
Matches 135;	Conservative 0;	Mismatches 66; Indels 0; Gaps 0
DQ	126	ACCACTTGATGTGTCGGTGTTCAGACACCCTTACTGAATCTGTTGCTCGGCTTCC 185
Db	209	ACCATGTTGTCTGGTGCCTGTTCAACAACCTCTTATGAACCCGCTGCCGCTTCA 150
DQ	186	TGGGGCTGCCTTAATCCTCATCAAGCCCGAGATCAAGAAGTGTGTGTAACCTGAAACCG 245
Db	149	TAGGTTTCACTTACTCTTCATGAAGTCTTAGAGCAAGAAATGATTGGCGAAGCTGACCGGG 90
DQ	246	CCGGCGCTTTTGGCTGCTCAAAGCCGAAGTCTCAACATCTGGCCGCGAATGGAAGCTGG 305
Db	89	CCGAGGCTAATGCTTCCACCGCCGAAGTGTCTGGAACATCTGGGCTTGACTTTGGGCAATCC 30
DQ	306	TGCCGCACTGTGCTCTCTGG 326
Db	29	TCATGACCATCTGCTCATCG 9
RESULT 7		
AKI3477/c		
ID	AKI3477	standard; DNA; 326 BP.
XX	AAKI3477;	
DT	05-NOV-2001	(first entry)
DE	Human brain expressed single exon probe SEQ ID NO: 13468.	
XX	Human; brain expressed exon; gene expression analysis; probe; microarray; Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer; ss.	
OS	Homo sapiens.	
XX	WO200157275-A2.	

PD 09-AUG-2001.
PF 30-JAN-2001; 2001WO-US000667.
PP
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
P1 Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
DR
XX
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
PS
PS Example 4; SEQ ID NO: 13468; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.

SQ	Sequence	326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
	Query Match	14.7%; Score 95.4; DB 22; Length 326;
	Best Local Similarity	67.2%; Pred. No. 5.8e-12;
	Matches 135; Conservative	0; Mismatches 66; Indels 0; Gaps 0
OY	126	ACCACTGATCTGGTCGGTGTTCAGCACCCCTCTAACCCTAAATCTGTGCTCGCCTGCC
Db	209	ACCATGTTGCTGGTCCCTGTTCAACAACCTCTTCATGAACCCCTGCCTCGGGCTTCA
OY	186	TGGGGTGCGCTACTCCATCAAAGCCCGAATCAAGAAGTGTGTGTACCTGGAACCGG
Db	149	TAGGGTTCACCTACTCTCATGAAGTCTTAGGGGACAGAAAGATGTTGGCACTGACCGGG
OY	246	CCCGCGCTTTTGCTCCAAAGCCAAGTCTCAACATCCTGAGCGCGAGTGTGACGCTGG
Db	89	CCGAGGCGTAATGCTCCACCGCCAAAGTGTCCAGAACATCTGGGCGCTGACTTTGGGCATCC
OY	306	TGCCGCACCTGCTGCTCCTGG 326
Db	29	TCATGACCATTTCTGCTCATCG 9

RESULT: 8
AAK39215/c
ID AAK39215 standard; DNA, 326 BP.
XX AAK39215;
AC
XX
DT 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 13772.
XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
OS Homo sapiens.
XX
XX MO200157276-A2.
XX
XX PN
XX PD 09-AUG-2001.


```
PF 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 13772; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX Sequence 326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
XX
XX Query Match 14.7%; Score 95.4; DB 22; Length 326;
XX Best Local Similarity 67.2%; Pred. No. 5.8e-12;
XX Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
XX
QY 126 ACCACTGATCTGCGTGGTTCAGACACCTTACCTGATCTGTGCTGGGCTTCC 185
DB 209 ACCATGTTGTCTGCTCCCTGTTCAACACCTCTTCAAGAACCTCTGCTGGGCTTCA 150
QY 186 TGGCGCTGACCTTACCTCATCAAGCCCGAGATCAAGAGTGTGTGACCTTGAAGCGG 245
DB 149 TAGCGTTACCTTACCTCATGAGTCTAAGGACAGAAAGATGTTGGCGACCTGACCGGG 90
QY 246 CCCGGCGTTTGGCTCTCAAGCCCAAGTGTCAACATCTCTGGCCGGAGTGTGACGCTGG 305
DB 89 CCCAGGCTATGCTCTCACCGCCAAAGTGCCTGAACATCTGGGCCCTGACTTTGGGCAATCC 30
QY 306 TGGCGCAGCTGCTGCTCTCTGG 326
DB 29 TCATGACCATTTCTGCTCATCG 9
XX
XX RESULT 9
XX AA120025/c
XX ID AA120025 standard; DNA; 326 BP.
XX
XX AA120025;
XX
XX 12-OCT-2001 (first entry)
XX
XX Probe #958 for gene expression analysis in human cervical cell sample.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; ss.
XX
XX Homo sapiens.
XX
XX MO200157278-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
XX
XX 04-FEB-2000; 2000US-0180312.
```

```
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 25; SEQ ID NO 9958; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENP). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
XX
XX Query Match 14.7%; Score 95.4; DB 22; Length 326;
XX Best Local Similarity 67.2%; Pred. No. 5.8e-12;
XX Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
XX
QY 126 ACCACTGATCTGCGTGGTTCAGACACCTTACCTGATCTGTGCTGGGCTTCC 185
DB 209 ACCATGTTGTCTGCTCCCTGTTCAACACCTCTTCAAGAACCTCTGCTGGGCTTCA 150
QY 186 TGGCGCTGACCTTACCTCATCAAGCCCGAGATCAAGAGTGTGTGACCTTGAAGCGG 245
DB 149 TAGCGTTACCTTACCTCATGAGTCTAAGGACAGAAAGATGTTGGCGACCTGACCGGG 90
QY 246 CCCGGCGTTTGGCTCTCAAGCCCAAGTGTCAACATCTCTGGCCGGAGTGTGACGCTGG 305
DB 89 CCCAGGCTATGCTCTCACCGCCAAAGTGCCTGAACATCTGGGCCCTGACTTTGGGCAATCC 30
QY 306 TGGCGCAGCTGCTGCTCTCTGG 326
DB 29 TCATGACCATTTCTGCTCATCG 9
XX
XX RESULT 10
XX AA145223/c
XX ID AA145223 standard; DNA; 326 BP.
XX
XX AA145223;
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #13909 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX MO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
XX
```

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024253.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID No 13909; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders.
XX
SQ Sequence 326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
Query Match 14.7%; Score 95.4; DB 22; Length 326;
Best Local Similarity 67.2%; Pred. No. 5.8e-12;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 126 ACCACTTGAATCTGCTGCTGCTTCAAGACCCCTTACTGTAATCTGCTGCTGCTTCC 185
DB 209 ACCATGTTCTGCTGCTGCTGCTTCAAGACCCCTTACTGTAATCTGCTGCTGCTTCA 150
QY 186 TGGCGTGGCTCTCTCATCAAGCCCGAGATCAGAAAGTGTTGATGACCTGGAAGCGG 245
DB 149 TAGCGTTACCTACTCTCATGAAGCTTAGGACAGAAAGATGTTGGACCTTACCGGGG 90
QY 246 CCGGCGGTTTGGCTCCAAAGCCAAAGTGTACAACTCTGCGCGGATGTGACGCTGG 305
DB 89 CCCAGGCTATGCTCCACCGCCAAAGTGTGAAACATCTGGGCCCTGACTTTGGGCATCC 30
QY 306 TGGCGCACTGCTGCTCTCTGG 326
DB 29 TCATGACCAATCTGCTCATCG 9
RESULT 11
AAI05732/C
ID AAI05732 standard; DNA; 326 BP.
XX
XX AAI05732;
XX
XX 09-OCT-2001 (first entry)
XX
XX DE Probe #5723 used to measure gene expression in human breast sample.
XX
XX KM Probe; human; breast disease; breast cancer; development disorder; ss;
XX KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX OS Homo sapiens.
XX
XX PN MO200157270-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 29-JAN-2001; 2001WO-US00661.
XX
XX 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024253.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
XX PT in a human breast -
XX
XX Claim 25; SEQ ID No 5723; 322bp; English.
XX
XX The present invention relates to novel single exon nucleic acid probes.
XX CC The present sequence is one such probe. The probes are useful for
XX CC measuring human gene expression in a human breast sample, where the probe
XX CC hybridizes at high stringency to a nucleic acid expressed in the human
XX CC breast. The probes are useful for predicting, diagnosing, grading,
XX CC staging, monitoring and prognosing diseases of the human breast,
XX CC particularly those diseases with polygenic aetiology. The diseases
XX CC include: breast cancer, disorders of development, inflammatory diseases
XX CC of the breast, fibrocystic changes, proliferative breast disease and
XX CC non-carcinoma tumours.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
Query Match 14.7%; Score 95.4; DB 22; Length 326;
Best Local Similarity 67.2%; Pred. No. 5.8e-12;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 126 ACCACTTGAATCTGCTGCTGCTTCAAGACCCCTTACTGTAATCTGCTGCTGCTTCC 185
DB 209 ACCATGTTCTGCTGCTGCTGCTTCAAGACCCCTTACTGTAATCTGCTGCTGCTTCA 150
QY 186 TGGCGTGGCTCTCTCATCAAGCCCGAGATCAGAAAGTGTTGATGACCTGGAAGCGG 245
DB 149 TAGCGTTACCTACTCTCATGAAGCTTAGGACAGAAAGATGTTGGACCTTACCGGGG 90
QY 246 CCGGCGGTTTGGCTCCAAAGCCAAAGTGTACAACTCTGCGCGGATGTGACGCTGG 305
DB 89 CCCAGGCTATGCTCCACCGCCAAAGTGTGAAACATCTGGGCCCTGACTTTGGGCATCC 30
QY 306 TGGCGCACTGCTGCTCTCTGG 326
DB 29 TCATGACCAATCTGCTCATCG 9
RESULT 12
ABS38800/C
ID ABS38800 standard; DNA; 326 BP.
XX
XX ABS38800;
XX
XX 25-FEB-2003 (first entry)
XX
XX DE Human liver single exon probe, SEQ ID No 13790.
XX
XX KM Human; single exon nucleic acid probe; liver; cirrhosis;
XX KM hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KM coronary heart disease; ss.
XX
XX OS Homo sapiens.
XX
XX PN MO200157273-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analysing gene expression in human adult liver -
XX
XX Claim 4; SEQ ID No 13790; 658bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult
XX liver. (1) may be used for predicting, measuring and displaying gene
XX expression in samples derived from human adult liver. The genes
XX identified may be involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which
XX is associated with coronary heart disease. ABS25011-ABS51005 represent
XX human liver single exon nucleic acid probes of the invention.
XX Note: The sequence information for this patent does not appear in the
XX printed specification but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
XX
XX Query Match 14.7%; Score 95.4; DB 23; Length 326;
XX Best Local Similarity 67.2%; Pred. No. 5.8e-12;
XX Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
XX
XX
XX 126 ACCACTGATCTGGTGGTTCAGACCTCTTACCTGATCTGTTGCTCCCTGCTTC 185
XX 209 ACCATGTTCTGCTGCTCCCTTCAACACCTCTTCAAGACCTGCTCCCTGCTTCA 150
XX 186 TGGCGCTGGCTTACCTCATCAAGCCGAGATCAGAGGTGGTGGTACCTGAAAGCG 245
XX 149 TAGCTTACCTACTCTCATGAGAGCTAGGAGACGAGAGTGGTGGCGACCTGACCGG 90
XX 246 CCCGCGCTTTGGCTCCCAAGCCAGATGCTACACATCTCTGGCCGCGATGTGACGCTGG 305
XX DB 89 CCCAGGCTATGCTCTCCACCGCAGATGCTGAAACATCTGGCCCTGATTTGGGATCC 30
XX QY 306 TGGCCGCACTGCTGCTCTCTGG 326
XX DB 29 TCATGACCATCTCTGCTCATCG 9
XX
XX
XX RESULT 13
XX ABS13299/c
XX ID ABS13299 standard; DNA; 326 BP.
XX
XX ABS13299;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon ORF from lung SEQ ID No 13290.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangiomyomatosis; Karsenger syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX
XX Claim 4; SEQ ID No 13290; 634bp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarray having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene
XX expression analysis, and for identifying exons in a gene, particularly
XX using human lung derived mRNA and for the study of lung diseases
XX such as asthma, lung cancer, chronic obstructive pulmonary disease
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
XX haemosiderosis, pulmonary histiocytosis, lymphangiomyomatosis,
XX pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
XX and hyaline membrane disease. The present sequence is a single exon
XX probe open reading frame of the invention.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX

Sequence 326 BP; 70 A; 79 C; 113 G; 64 T; 0 other;
Query Match 14.7%; Score 95.4; DB 24; Length 326;
Best Local Similarity 67.2%; Pred. No. 5.8e-12;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 126 ACCACTTGATCTGGTGGTGTTCAGACCCCTCTGACTGAATCTGTGTGCTGCTGCTTCC 185
DB 209 ACCATGTTGCTGGGCTCCCTGTCAACCCCTTCATGAACCCCTGCTGCTGGGCTTCA 150
QY 186 TGGGCTGGCCCTACTCTCAATCAAGCCCGAGATCAAGATGTGTGTGACCTGGAAGCGG 245
DB 149 TAGGTTACCTTACTCCATGAAGTCTAGGACAGGAAGATGTGTGGCGACTGACCGGGG 90
QY 246 CCCGCGCTTTTGGCTCCAAAGCCAAAGTCTACCAATCTCTGCGCGCATGTGAACGCTGG 305
DB 89 CCCAGGCTATGCTCTCCACCGCCCAAGTCCCTGAACATCTGGGCTCTGACTTTGGGCACTCC 30
QY 306 TGGCGCCACTGCTGCTCTCTGG 326
DB 29 TCATGACCATCTGCTCATCG 9
RESULT 14
ABA42012/c
ID ABA42012 standard; DNA; 459 BP.
XX ABA42012;
XX ABA42012;
XX 01-FEB-2002 (first entry)
XX Human breast cell single exon nucleic acid probe #707.
XX Human; microarray; single exon probe; gene expression; breast;
XX disease; cancer; ss.
XX Homo sapiens.
XX WO200157271-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00662.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes
XX Claim 1; SEQ ID NO 707; 327pp + sequence listing; English.
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing breast disease. Gene

expression analysis is useful for assessing the toxicity of chemical
agents on cells. The microarray of this invention presents a far greater
diversity of probes for measuring gene expression, with far less bias
than expressed sequence tag microarrays. The method is suitable for
rapid production of functional information from genomic sequence. The
present sequence is a single exon nucleic acid probe of the invention.
Note: The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at http://wipo.int/pub/published_pot_sequences.
Sequence 459 BP; 115 A; 96 C; 154 G; 94 T; 0 other;
Query Match 14.7%; Score 95.4; DB 22; Length 459;
Best Local Similarity 67.2%; Pred. No. 6.1e-12;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 126 ACCACTTGATCTGGTGGTGTTCAGACCCCTCTGACTGAATCTGTGTGCTGCTGCTTCC 185
DB 406 ACCATGTTGCTGGGCTCCCTGTCAACCCCTTCATGAACCCCTGCTGCTGGGCTTCA 347
QY 186 TGGGCTGGCCCTACTCTCAATCAAGCCCGAGATCAAGATGTGTGTGACCTGGAAGCGG 245
DB 346 TAGGTTACCTTACTCCATGAAGTCTAGGACAGGAAGATGTGTGGCGACTGACCGGGG 287
QY 246 CCCGCGCTTTTGGCTCCAAAGCCAAAGTCTACCAATCTCTGCGCGCATGTGAACGCTGG 305
DB 286 CCCAGGCTATGCTCTCCACCGCCCAAGTCCCTGAACATCTGGGCTCTGACTTTGGGCACTCC 227
QY 306 TGGCGCCACTGCTGCTCTCTGG 326
DB 226 TCATGACCATCTGCTCATCG 206
RESULT 15
ABA52433/c
ID ABA52433 standard; DNA; 459 BP.
XX ABA52433;
XX ABA52433;
XX 01-FEB-2002 (first entry)
XX Human foetal liver single exon nucleic acid probe #738.
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX Homo sapiens.
XX WO200157277-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00669.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-493447/52.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human foetal liver -
XX Claim 1; SEQ ID NO 738; 639pp + sequence listing; English.
XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 459 BP; 115 A; 96 C; 154 G; 94 T; 0 other;

Query Match 14.7%; Score 95.4; DB 22; Length 459;
Best Local Similarity 67.2%; Fred. No. 6.1e-12;

Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy	126	ACCACTTGATCTGGTGGGTTTCAAGACCCCTCTACCTGAATCTGTGCTCGGCTTCC	185
Db	406	ACCATGTTGTGTGGTCCCTGTTTCAACACCTCTTCATGAAACCCCTGCTGGGCTTCA	347
Qy	186	TGGCGCTGGCCTTACTCCATCAAGGCCGAGATCAGAAAGTGTGTGACCTGGAAGCGG	245
Db	346	TAGCGTTCACCTACTCCATGAAGTCTAGGACAGGAAGATGTGGCGACCTGACCGGGG	287
Qy	246	CCCGGCGTTTGGCTCCAAAGCTGACATCTGCGCCGATGTGAAGCTGG	305
Db	286	CCGAGGCTATGCTCCACCGCAAGTGCCTGAACATCTGGGCCCTGACTTGGGCATCC	227
Qy	306	TGCCGCCACTGTGCTGCTCGG	326
Db	226	TCATGACCATTTGCTCATCG	206

Search completed: January 31, 2004, 13:43:02

Job time : 290 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2004, 13:23:10 ; Search time 2842 Seconds
(without alignments)
9370.924 Million cell updates/sec

Title: US-09-914-815A-10

Perfect score: 651
Sequence: 1 ggaagagacggcgctggaac.....cccgggccctaactctgcc 651

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBankl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_srb: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_srs: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rod: *
36: em_hcg_mam: *
37: em_hcg_vtl: *
38: em_sy: *
39: em_hcgo_hum: *
40: em_hcgo_mus: *
41: em_hcgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	446.6	68.6	95038	9 AF015416	AF015416 Homo sapi
C 2	446.6	68.6	143835	9 AC136475	AC136475 Homo sapi
C 3	446.6	68.6	156819	2 AP006283	AP006283 Homo sapi
C 4	446.6	68.6	175416	2 AP006286	AP006286 Homo sapi
C 5	446.6	68.6	187160	9 AC138230	AC138230 Homo sapi
C 6	284.6	43.7	499	10 MMU9781	AC1009781 Mus muscu
C 7	173	26.6	238533	2 AC132968	AC132968 Rattus no
C 8	167.6	25.7	200734	2 AC109272	AC109272 Mus muscu
C 9	167.6	25.7	251206	2 AC107815	AC107815 Mus muscu
C 10	95.4	14.7	136098	9 AC006970	AC006970 Homo sapi
C 11	93.8	14.4	51891	2 AC074103	AC074103 Homo sapi
C 12	93.4	14.3	678	9 BC009696	BC009696 Homo sapi
C 13	93.4	14.3	905	6 AX337050	AX337050 Sequence
C 14	93.4	14.3	905	9 HS18D	X57351 Human 1-8D
C 15	93.2	14.3	402	9 BT006892	BT006892 Homo sapi
C 16	93.2	14.3	402	12 BT007876	BT007876 Synthetic
C 17	93.2	14.3	637	9 BC006794	BC006794 Homo sapi
C 18	93.2	14.3	642	6 AR225472	AR225472 Sequence
C 19	93.2	14.3	642	6 AX321588	AX321588 Sequence
C 20	93.2	14.3	645	9 BC022439	BC022439 Homo sapi
C 21	93.2	14.3	695	6 AX281855	AX281855 Sequence
C 22	93.2	14.3	749	9 BC008417	BC008417 Homo sapi
C 23	93.2	14.3	808	9 HS18U	X57352 Human 1-8U
C 24	91.6	14.1	37918	9 AP005232	AP005232 Homo sapi
C 25	91.6	14.1	222765	2 AC144988	AC144988 Gorilla g
C 26	91.2	14.0	378	9 BT0007173	BT0007173 Homo sapi
C 27	91.2	14.0	378	12 BT007946	BT007946 Synthetic
C 28	91.2	14.0	647	9 HSRNA327	X84958 H. sapiens m
C 29	91.2	14.0	683	9 BC000897	BC000897 Homo sapi
C 30	91.2	14.0	842	6 AX397528	AX397528 Sequence
C 31	91.2	14.0	851	6 AX014337	AX014337 Sequence
C 32	91.2	14.0	853	6 AX302539	AX302539 Sequence
C 33	91.2	14.0	853	6 AX409461	AX409461 Sequence
C 34	91.2	14.0	853	9 HM927A	U04164 Human inter
C 35	91.2	14.0	1095	6 AX281854	AX281854 Sequence
C 36	91.2	14.0	245	6 AX340862	AX340862 Sequence
C 37	90.8	13.9	579	4 AF272041	AF272041 Bos tauru
C 38	90.8	13.9	107717	9 AC137055	AC137055 Homo sapi
C 39	90.8	13.9	110000	2 AC004085_1	Continuation (2 of
C 40	90.8	13.9	110000	2 AC004085_2	Continuation (3 of
C 41	90.2	13.9	122351	9 AC023157	AC023157 Homo sapi
C 42	89.6	13.8	905	10 BC027285	BC027285 Mus muscu
C 43	89.6	13.8	84979	9 AC116049	AC116049 Homo sapi
C 44	89.6	13.8	176104	2 AC016193	AC016193 Homo sapi
C 45	89.6	13.8	189572	2 AC084812	AC084812 Homo sapi

ALIGNMENTS

RESULT 1
AF015416/c 95038 bp DNA linear PRI 20-AUG-1997
LOCUS Homo sapiens chromosome 11 from 11p15.5 region, complete sequence.
DEFINITION AF015416
ACCESSION AF015416.1 GI:2335202
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 95038)
Evans, G.A., Bradbury, P., Brignac, S., Bunelster, R., Davis, J.,
Davies, C.J., Davis, C., English, C., Fondon, T., Franklin, T.L.,
Garner, H.R., Gordon, M., Gotway, G., Grant, O., Hahner, L., Harris, J.,

Pred. No. is the number of results predicted by chance to have a

TITLE
JOURNAL
REFERENCE
AUTHORS
Hinson, S., Megarity, C., Narayanaswamy, U., Newton, J., O'Brien, K.,
Olliver, T., Patel, P., Prober, S., Rayner, S., Schageman, J.,
Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T.,
Wilson, R. and Burbee, D.
95 Kb from Four Overlapping Human Chromosome 11p15.5 Cosmids
Unpublished
2 (bases 1 to 95038)
Evans, G.A., Bradbury, P., Brignac, S., Bumeister, R., Davie, J.,
Davies, C.J., Davis, C., English, C., Fondon, T., Franklin, T.L.,
Gainer, H.R., Gordon, M., Grant, O., Hahner, L., Harris, J.,
Hinson, S., Megarity, C., Narayanaswamy, U., Newton, J., O'Brien, K.,
Olliver, T., Patel, P., Prober, S., Rayner, S., Schageman, J.,
Schilling, P., Schultz, R., Syed, M., Valenzuela, D., Ward, T.,
Wilson, R. and Burbee, D.
Direct Submission
Submitted (18-JUL-1997) Genome Science and Technology Center,
University of Texas Southwestern Medical Center at Dallas, 5323
Harry Hines Blvd, Dallas, TX 75235-8591, USA
FEATURES
SOURCE
Location/Qualifiers
1..95038
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11p15.5"
711..975
/rpt_family="Alu"
repeat_region
2728..3022
/rpt_family="Alu"
repeat_region
3363..3719
/rpt_family="THE1"
repeat_region
5634..5909
/rpt_family="Alu"
repeat_region
complement(6720..6959)
/rpt_family="Alu"
repeat_region
10560..10839
complement(11832..12455)
/rpt_family="Alu"
repeat_region
12921..13405
/rpt_family="Alu"
repeat_region
13718..14314
/rpt_family="L1"
repeat_region
14617..14779
/rpt_family="Alu"
repeat_region
15396..15695
/rpt_family="Alu"
repeat_region
15722..15994
/rpt_family="Alu"
repeat_region
complement(16696..16997)
/rpt_family="Alu"
repeat_region
complement(18996..19286)
/rpt_family="Alu"
repeat_region
complement(19458..19747)
/rpt_family="Alu"
repeat_region
complement(19777..20045)
/rpt_family="Alu"
repeat_region
complement(22885..23002)
/rpt_family="MIR"
repeat_region
23109..23361
/rpt_family="Alu"
repeat_region
23292..23481
/rpt_family="MER46"
repeat_region
complement(23877..23944)
/rpt_family="MIR"
repeat_region
complement(24695..24986)
/rpt_family="Alu"
repeat_region
complement(25429..26012)
/rpt_family="Alu"
repeat_region
complement(27380..27676)
/rpt_family="Alu"
repeat_region
complement(29422..29486)
/rpt_family="MER42"
repeat_region
complement(29514..29802)
/rpt_family="Alu"
repeat_region
30797..31075
/rpt_family="Alu"
repeat_region
31688..31792
/rpt_family="Alu"
repeat_region
32125..32412
/rpt_family="Alu"
repeat_region
complement(38871..39186)
/rpt_family="Alu"
repeat_region
42952..43094
/rpt_family="MERS"
repeat_region
43106..43388
/rpt_family="Alu"
repeat_region
43754..44216
/rpt_family="MLT2B2"
repeat_region
complement(44815..45136)
/rpt_family="MER42"
repeat_region
45145..45484
/rpt_family="Alu"
repeat_region
45740..45995
/rpt_family="Alu"
repeat_region
46375..46646
/rpt_family="Alu"
repeat_region
47145..47738
/rpt_family="Alu"
repeat_region
48353..48626
/rpt_family="Alu"
repeat_region
49567..49842
/rpt_family="Alu"
repeat_region
49940..50211
/rpt_family="Alu"
repeat_region
complement(50615..50907)
/rpt_family="Alu"
repeat_region
complement(53001..53290)
/rpt_family="Alu"
repeat_region
complement(54161..54512)
/rpt_family="L1"
repeat_region
54533..54823
/rpt_family="Alu"
repeat_region
57899..58113
/rpt_family="Alu"
repeat_region
complement(58795..59044)
/rpt_family="Alu"
repeat_region
complement(59414..60023)
/rpt_family="L1"
repeat_region
complement(60177..60311)
/rpt_family="Alu"
repeat_region
complement(60770..61108)
/rpt_family="L1"
repeat_region
complement(61207..61557)
/rpt_family="THE1"
repeat_region
61670..61942
/rpt_family="Alu"
repeat_region
61944..62436
/rpt_family="MER41"
repeat_region
complement(67807..68105)
/rpt_family="Alu"
repeat_region
70961..71049
/rpt_family="Alu"
repeat_region
71224..71523
/rpt_family="Alu"
repeat_region
complement(85033..85397)
/rpt_family="MER4"
repeat_region
complement(85414..85620)
/rpt_family="Alu"
repeat_region
complement(85701..85804)
/rpt_family="MER4"
repeat_region
complement(85895..86028)
/rpt_family="MER4"
repeat_region
86180..86477
/rpt_family="Alu"
repeat_region
86497..86780

Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L28535
 Center clone name: 326_C3

Only the first 143.8 kilobases of this clone are being submitted.
 The remainder overlaps accession number AC138230 [WIGR project
 L29013].

FEATURES
 SOURCE

Location/Qualifiers
 1. 143835
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-326C3"
 /clone_1lb="RPC1-11 Human Male BAC"
 1712..2006
 /rpt_family="AluSg"
 2347..2703
 /rpt_family="TH31B"
 4157..4302
 /rpt_family="L1MB4a"
 4435..4455
 /rpt_family="AT_rich"
 4528..4591
 /rpt_family="AT_rich"
 4592..4901
 /rpt_family="AluSx"
 4902..4968
 /rpt_family="AT_rich"
 4979..5091
 /rpt_family="L1MB4a"
 5154..5251
 /rpt_family="L1MB4a"
 5620..5683
 /rpt_family="L1MB4a"
 complement(5684..5954)
 /rpt_family="AluJo"
 5955..6050
 /rpt_family="L1MB4a"
 complement(6059..6206)
 /rpt_family="MER46C"
 6349..6578
 /rpt_family="L1MB4a"
 6792..7203
 /rpt_family="L2"
 complement(8884..9052)
 /rpt_family="L1MCS"
 9225..9287
 /rpt_family="MER68-Int"
 9676..9986
 /rpt_family="AluSg"
 9994..10032
 /rpt_family="AT_rich"
 10182..10417
 /rpt_family="L1MB4a"
 complement(10418..10946)
 /rpt_family="L1R59"
 complement(10947..11251)
 /rpt_family="AluSx"
 complement(11252..11272)
 /rpt_family="L1R59"
 complement(11290..11595)
 /rpt_family="AluSg"
 11702..12014
 /rpt_family="AluSg"
 12047..12137
 /rpt_family="L1M4"
 complement(12159..12465)

/rpt_family="AluSx"
 12466..12605
 /rpt_family="L2"
 12763..12832
 /rpt_family="L2"
 12898..13715
 /rpt_family="L1MB5"
 13716..14007
 /rpt_family="AluSc"
 14018..14320
 /rpt_family="AluSx"
 14321..14603
 /rpt_family="L1MB5"
 14611..14772
 /rpt_family="AluSg/x"
 15386..15685
 /rpt_family="AluSx"
 15701..15984
 /rpt_family="AluJo"
 complement(16654..16949)
 /rpt_family="AluSx"
 17027..17095
 /rpt_family="L1MCS"
 complement(18948..19258)
 /rpt_family="AluY"
 19299..19409
 /rpt_family="L2"
 complement(19426..19719)
 /rpt_family="AluSg"
 complement(19742..20040)
 /rpt_family="AluJo"
 20207..20435
 /rpt_family="L1MCS/D"
 20906..21032
 /rpt_family="CCCCCG)n"
 complement(21181..21341)
 /rpt_family="MIR"
 complement(22663..22828)

Query Match 68.6%; Score 446.6; DB 9; Length 143835;

Best Local Similarity 99.1%; Pred. No. 2.3e-78; Mismatches 4; Indels 0; Gaps 0;

Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 199 CTCATCAAGCCGAGATCAGAGGTGTGTGACTTGAAGCCGCCGCGTTTGG 258
 DB 83267 CTCTGACAGCCGAGATCAGAGGTGTGTGACTTGAAGCCGCCGCGTTTGG 83208
 QY 259 CTCGAAGCCAGGTCTACACATCTGCGCGGATGTGACGCTGTCGCCACTGCT 318
 DB 83207 CTCGAAGCCAGGTCTACACATCTGCGCGGATGTGACGCTGTCGCCACTGCT 83148
 QY 319 GCTCCGAGGCTGTGTGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 378
 DB 83147 GCTCCGAGGCTGTGTGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 83088
 QY 379 CGCTCTTTCAGACCAAGTTTGTATGACGCGACTATGACTGACAGGCTGGTCTGATC 438
 DB 83087 CGCTCTTTCAGACCAAGTTTGTATGACGCGACTATGACTGACAGGCTGGTCTGATC 83028
 QY 439 TGGGAGCAAGCCGAGACCTGACCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 498
 DB 83027 TGGGAGCAAGCCGAGACCTGACCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 82968
 QY 499 CCGGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 558
 DB 82967 CCGGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 82908
 QY 559 CCTTCAGCCCAATGAGGCACTTAAGGCTGAACAGTCAAGCCCGGGGTCTTCAACC 618
 DB 82807 CCTTCAGCCCAATGAGGCACTTAAGGCTGAACAGTCAAGCCCGGGGTCTTCAACC 82848
 QY 619 TAACCGAGAGTTCCGGGCGCTTAATCTTCC 651


```
misc_feature 125282..129184
              /note="assembly_fragment clone_end:T7 vector_side:right"
misc_feature 129285..137260
              /note="assembly_fragment"
misc_feature 137361..143542
              /note="assembly_fragment"
misc_feature 143643..146891
              /note="assembly_fragment"
misc_feature 146992..149568
              /note="assembly_fragment"
misc_feature 149669..152678
              /note="assembly_fragment"
misc_feature 152779..154135
              /note="assembly_fragment"
misc_feature 154236..155662
              /note="assembly_fragment"
misc_feature 155763..156819
              /note="assembly_fragment"
BASE COUNT 37089 a 39987 c 40186 g 37757 t 1800 others
ORIGIN
Query Match 68.6%; Score 446.6; DB 2; Length 156819;
            Best Local Similarity 99.1%; Pred. No. 2.3e-78;
Matches 449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 199 CTCATCAAGCCCGAGATCGAAGGTGTGTGACCTGGAAGCGCCGCGTTTGG 258
DB 92819 CTCCTGAGGCGCGAGATCGAAGGTGTGTGACCTGGAAGCGCGCGTTTGG 92878
QY 259 CTCGAAGCCGAAGTCTCAACATCTGCGCGAGTGTGAGCGTGGTCCGCAATGCT 318
DB 92879 CTCGAAGCCGAAGTCTCAACATCTGCGCGAGTGTGAGCGTGGTCCGCAATGCT 92938
QY 319 GCTCTGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 378
DB 92939 GCTCTGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 92998
QY 379 CGCTTCTTCAGACCAAGTTTGTATGACGCGACTATATGACGAGGTGGTCTGATC 438
DB 92999 CGCTTCTTCAGACCAAGTTTGTATGACGCGACTATATGACGAGGTGGTCTGATC 93058
QY 439 TGGGGCACTAGACCCCAAGACACTGACCCAGGCTGTGCGCCCTGGGGCCCAATCTGACT 498
DB 93059 TGGGGCACTAGACCCCAAGACACTGACCCAGGCTGTGCGCCCTGGGGCCCAATCTGACT 93118
QY 499 CCGCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 558
DB 93119 CCGCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 93178
QY 559 CCGTTCAGACCCCAAGACACTGAGGCTGAACCAAGTCAAGCCCGGGGTCTTCAACC 618
DB 93179 CCGTTCAGACCCCAAGACACTGAGGCTGAACCAAGTCAAGCCCGGGGTCTTCAACC 93238
QY 619 TAACCGAGAGTTCGGGGCCCTAATCTGACC 651
DB 93239 TAACCGAGAGTTCGGGGCCCTAATCTGACC 93271

RESULT 4
AP006286/c 175416 bp DNA linear HTG 26-MAR-2003
LOCUS Homo sapiens chromosome 11 clone RP13-25N22 map 11p, WORKING DRAFT
DEFINITION AP006286
ACCESSION AP006286.1 GI:29243346
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Toroki,Y., Watanabe,H. and Sakaki,Y.
```

```
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Homo sapiens genomic DNA of 11p
Published Only in Database (2003)
2 (bases 1 to 175416)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Toroki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (24-MAR-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Hmndraft11
Center clone name: RP13-25N22
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 16074 bases at least Q40
Consensus quality: 167722 bases at least Q30
Consensus quality: 171662 bases at least Q20
Insert size: 173316; sum-of-coverage
Quality coverage: 7.92x in Q20 bases; sum-of-coverage
-----
NOTE: This is a 'working draft' sequence. It currently consists of
22 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 33331 contig of 33331 bp in length
33232 54098 contig of 20867 bp in length
54199 72037 contig of 17839 bp in length
72138 83277 contig of 11140 bp in length
83378 98599 contig of 15222 bp in length
98700 105513 contig of 16814 bp in length
105614 116229 contig of 10616 bp in length
116330 121964 contig of 5635 bp in length
122065 127627 contig of 5563 bp in length
131971 contig of 4444 bp in length
137506 contig of 5435 bp in length
143829 contig of 6223 bp in length
143930 148539 contig of 4610 bp in length
148640 153148 contig of 4609 bp in length
153249 157509 contig of 4261 bp in length
157610 160561 contig of 2952 bp in length
160662 164778 contig of 4117 bp in length
164879 168269 contig of 3391 bp in length
168370 170689 contig of 2220 bp in length
170790 172833 contig of 2044 bp in length
172934 174155 contig of 1222 bp in length
174256 175416 contig of 1161 bp in length.
NOTE: This is a 'working draft' sequence. It currently
consists of 22 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 33331: contig of 33331 bp in length
* 33332 gap of 100 bp
* 33332 33231: contig of 20867 bp in length
* 54099 54198: gap of 100 bp
* 54199 72037: contig of 17839 bp in length
* 72038 72137: gap of 100 bp
* 72138 83277: contig of 11140 bp in length
```


Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fark, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatae, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhngang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (20-DEC-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187160)

REFERENCE
AUTHORS
Birtren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, D., Fark, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatae, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, T., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhngang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (10-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 187160)

REFERENCE
AUTHORS
Birtren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, D., Fark, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatae, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, T., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhngang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-MAY-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 22, 2003 this sequence version replaced g1:29423936.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L29013
Center clone name: 317_D_12

Location/Qualifiers
1. 187160
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP13-317D12"
/clone_1kb="RP13-13 Human Female BAC"
repeat_region
135..700
/rpt_family="L2"
repeat_region
1134..1310
/rpt_family="MIR"
complement(1311..1614)
/rpt_family="AluY"
repeat_region
1615..1638
/rpt_family="MIR"
repeat_region
1849..1938
/rpt_family="L1ME"
complement(4297..4385)
/rpt_family="L2"
repeat_region
4463..4561
/rpt_family="FRAM"
complement(4573..4613)
/rpt_family="MAD1"
4733..5040
/rpt_family="AluY"
complement(5094..5383)
/rpt_family="L2"
complement(5379..5505)
/rpt_family="L2"
complement(5501..5627)
/rpt_family="L2"
complement(5667..6096)
/rpt_family="L2"
repeat_region
6542..6666
/rpt_family="MER41E"
6639..6719
/rpt_family="GC-rich"
6917..6996
/rpt_family="C-rich"
7136..7329
/rpt_family="C-rich"
15461..15483
/rpt_family="(TG)n"
18178..18413
/rpt_family="L2"
complement(18543..18907)
/rpt_family="MER4A"
complement(18913..19138)
/rpt_family="AluSg/x"
complement(19139..19318)
/rpt_family="MER4D"
complement(19319..19688)
/rpt_family="MER4E1"
19689..19992
/rpt_family="AluSg"
19994..20308
/rpt_family="AluSg"
complement(20309..20630)
/rpt_family="MER4E1"
20632..20763
/rpt_family="AluSg/x"
complement(20785..20882)
/rpt_family="MER4E"
complement(20886..21091)
repeat_region

Db 249 CTGGAGAGCTGCAAGGACGATGAGCTCCAAAGCAAGTCTCAACAACATCTGCTGCAT 308
 Qy 295 GTGACGCTGTGCTGCGGACACTGTCTCTGCGGCTGTGAGTGTGACTGATGCTGCACCT 354
 Db 309 GTGACATGTGGTGTGCCCCATTGCTGCTCTGCGACTGTGTGTGACTGTGCGCTTTGACCT 368
 Qy 355 GGGCCGCTGGCCCAAGACTGTGCGCTTTCTTCAAGCAAGTTTATGATACCGGACTA 414
 Db 369 GTCCAAATTGACCAAAACTCTGCGGCTTTCTTCAAGCAAGTTTATGATGAGGACTA 428
 Qy 415 TGACTGACAGCTGGTCTG 435
 Db 429 TAACTAAGATGTTCCGACCTG 449

RESULT 7
 AC132968/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-172C3, *** SEQUENCING IN PROGRESS
 ACCESSION AC132968 238533 bp DNA linear HTG 20-NOV-2002
 VERSION AC132968.4 GI:25139219
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 AUTHORS 1 (bases 1 to 238533)
 Mizny, D., Marie, Metzger, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Anguiano, D., Anyalbechli, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, K., Barnstead, M., Benahmed, F., Bialwalo, K., Blair, D., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, K., Cavazos, I., Ceasar, H., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Calderon, E., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, R., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falle, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Gebreyes, E., Geer, K., Gill, R., Grady, W., Guerra, R., Guavara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jollivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, Y., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loulsegged, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Nangum, A., Mangun, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Miliostavlevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nsamenang, A., Nsamenang, O., Okunnu, G., Olariunso, A., Pal, S., Parks, K., Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Plazio, B., Quiroz, J., Rachin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Sytek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,

REFERENCE
 AUTHORS 1 (bases 1 to 238533)
 Williams, G., Wilson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 238533)
 Rat Genome Sequencing Consortium.
 Rat Genome Sequencing Consortium.
 Submitted (05-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 238533)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center

Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: XEMS
 Center clone name: CH230-172C3

Summary Statistics

Assembly program: Phrap; version 0.990329
 Consensus quality: 223317 bases at least Q40
 Consensus quality: 225463 bases at least Q30
 Consensus quality: 226555 bases at least Q20
 Estimated insert size: 230654; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draat_data.html).
 NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 222461: contig of 222461 bp in length
 * 222462 222561: gap of unknown length
 * 222562 224949: contig of 2388 bp in length
 * 224950 225049: gap of unknown length
 * 225050 228242: contig of 3193 bp in length
 * 228243 228342: gap of unknown length
 * 228343 228533: contig of 10191 bp in length.

Location/Qualifiers

FEATURES

1. 238533
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"

misc_feature

1..2349
 /note="wgs_contig"


```

TITLE
JOURNAL
COMMENT

Submitted (06-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 6, 2003 this sequence version replaced gi:20153115.
All repeats were identified using RepeatMasker:
Shult, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/BM/RepeatMasker.html

-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L20545
Center clone name: 114_A_6
-----
Summary Statistics
Sequencing vector: Plasmid/n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 250046 bases at least Q40
Consensus quality: 250590 bases at least Q30
Consensus quality: 250806 bases at least Q20
Insert size: 230000; agarose-FP
Insert size: 250906; sum-of-contigs
Quality coverage: 12.6 in Q20 bases; agarose-FP
Quality coverage: 11.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 29459: contig of 29459 bp in length
*      29460 29559: gap of 100 bp
*      29560 30528: contig of 965 bp in length
*      30529 30628: gap of 100 bp
*      30629 142773: contig of 112145 bp in length
*      142774 142873: gap of 100 bp
*      142874 251206: contig of 108333 bp in length.
*
Location/Qualifiers
1..251206
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone_name="RP23-114A6"
/clone_lib="RPc1-23 Female Mouse BAC"
1..29459
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
29560..30528
/note="assembly_fragment"
30629..142773
/note="assembly_fragment"
142874..251206
/note="assembly_fragment"
clone_end:T7
vector_side:right"

```

Query Match	25.7%	Score 167.6	DB 2	Length 251206
Best Local Similarity	63.8%	Pred. No. 1.6e-23		
Matches 277	Conservative	0	Mismatches 144	Indels 13
				Gaps 1
QY	201	CCATCAAGAGCCCGAAGATCAAGAGTGGTGTGACCTGGAAAGCGGCCCGGCTTTGGCT	260	
Db	131688	CCCTGACGGCCCGAGACCAAGATGCTGGAACTTGGAGGCTGCAGAGGAGTATGCT	13168	
QY	261	CCAAAGCCAAGTGTACCAATCCTGTGCGCGCATGTGAGACGCTGTGCTCCGCACTGTGC	320	
Db	131628	CCAAAGCCAAGTGTACCAATCCTGTGCGCGCATGTGAGACGCTGTGCTCCGCACTGTGC	13155	
QY	321	TCCTGGGGCTGTGTGTGACTGTGTGCTGCTGCACTGCGCCGGCTGGCCAAAGACTCTGCGG	380	
Db	131568	TCCTGGGAGCTGTGTGTGACTGTGTGCTGCTGCACTGCGCCGGCTGGCCAAAGACTCTGCGG	13156	
QY	381	CCCTCTTACACCAAGATTGATGACCGCGGACATATGACTACAGGCTGGGTCTGATCTG	440	
Db	131508	CTTCTTACACCAAGATTGATGAGAGGACATATGACTACAGGCTGGGTCTGATCTG	13144	
QY	441	GGGACACTAGAGCCCAAGACACTGACCCCGAGGCTGTGCGCCCTGGGGCCCAATACTGACTCC	500	
Db	131453	-----TCCCTGAAACCGAGAGACACATGTGACGTGTGCTGTGCGCCCAACAGCTCC	13144	
QY	501	CCGAGAGCTGAGCCCTCTCTGTGTGGGGCTTCATCCCTGCCCCATCCTGATCTGGAGCC	560	
Db	131401	TGGAGATGTACAGCTACTATGACCTGATCCCTGCGGAGCTGTGAGAGGCGCG	13133	
QY	561	CTCCAGCCCAATATGGGACCTTAAGCTGATACATGACACACCCCGGGGTCTTACCTTA	620	
Db	131341	CTCTTGTGTCCAGATGTGTGTGTGATGCTGCTGCCACATCAATCTCCCTGACATTAATTAT	13122	
QY	621	ACCCGAGAGTTCCC	634	
Db	131281	AACTTAGAGGATCC	131288	
RESULT 10				
LOCUS	AC006970	136098 bp	DNA	linear PRI 27-SEP-2000
DEFINITION	Homo sapiens clone RP4-725G10, complete sequence.			
ACCESSION	AC006970			
VERSION	AC006970.6	GI:10312289		
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	1 (bases 1 to 136098)			
JOURNAL	Waterston, R.H.			
REFERENCE	The sequence of Homo sapiens clone			
AUTHORS	Unpublished			
TITLE	2 (bases 1 to 136098)			
JOURNAL	Waterston, R.H.			
REFERENCE	Submitted (05-MAR-1999) Genome Sequencing Center, Washington			
AUTHORS	University School of Medicine, 444 Forest Park Parkway, St. Louis,			
TITLE	MO 63108, USA			
JOURNAL	3 (bases 1 to 136098)			
REFERENCE	Waterston, R.H.			
AUTHORS	Direct Submission			
TITLE	Submitted (27-SEP-2000) Genome Sequencing Center, Washington			
JOURNAL	University School of Medicine, 444 Forest Park Parkway, St. Louis,			
COMMENT	MO 63108, USA			
FEATURES	On Sep 27, 2000 this sequence version replaced gi:9838025.			
SOURCE	Center Project name: H DU0725G10.			
	Location/Qualifiers			
	1. 136098			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			

```

/cclone="RP4-725G10"
BASE COUNT      38294 a 29779 c 30262 g 37763 t
ORIGIN
Query Match      14.7%; Score 95.4; DB 9; Length 136098;
Best Local Similarity 67.2%; Pred. No. 2.8e-09;
Matches 135; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 126 ACCACTGATCTGGTGGTGTGAGACACCTTACCTGATCTGTGCTGCTGCTTC 185
DB 68841 ACCATGTTGCTGGTCCCTGTTCAACACCTTTCATGAACCCCTGCTGCTGCTTCA 68900
QY 186 TGGGGCTGAGCTTCTTCAATCAAGCCGAGATCAAGAGTGTGTGACTGGAAGCG 245
DB 68901 TAGGTTTCACTTCTTCAATGAAGTCTAGAGGACAGAAATGTTGGCGACTGACCGGG 68960
QY 246 CCCGCGCTTTGGCTCCAAAGCAAGTCTCAACATCTTGGCCGAGTGAAGCGCTGG 305
DB 68961 CCCAGGCTATGCTCTCCACCGCCAGTCTGAAACATCTGGGCTTGACTTTGGGCACTCC 69020
QY 306 TGGCGGCACCTGCTGCTCTGG 326
DB 69021 TCATGACCATCTGCTCATCG 69041

```

```

RESULT 11
AC074103 51891 bp DNA linear HTG 22-NOV-2000
LOCUS Homo sapiens chromosome UNK clone RP11-782B10, WORKING DRAFT
DEFINITION AC074103
SEQUENCE 3 unordered pieces.
AC074103
AC074103.4 GI:11276323
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 51891)
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 51891)
Waterston, R.H.
Direct Submission
Submitted (14-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Nov 22, 2000 this sequence version replaced gi:9954869.
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H.NH0782B10
----- Summary Statistics -----
Sequencing vector: M13; 95%
Sequencing vector: Plasmid; 5%
Chemistry: Dye-terminator Big Dye; 5% of reads
Chemistry: Dye-terminator Big Dye; 5% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: bases at least Q40
Consensus quality: bases at least Q30
Consensus quality: bases at least Q20
Insert size: 19200; agarose-fp
Insert size: 179562; sum-of-ctnigs
Quality coverage: 4.57 in Q20 bases; agarose-fp
Quality coverage: 4.77 in Q20 bases; sum-of-ctnigs

```

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2923: contig of 2923 bp in length
2924 3023: gap of unknown length
3024 15691: contig of 12668 bp in length
15692 15791: gap of unknown length
15792 51891: contig of 36100 bp in length.
location/Qualifiers
1. 51891
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="UNK"
/clone="RP11-782B10"
1. 2923
/note="assembly_name:Contig23"
3024. 15691
/note="assembly_name:Contig34"
15792. 51891
/note="assembly_name:Contig37"
BASE COUNT 13545 a 12146 c 12003 g 13995 t 202 others
ORIGIN
Query Match      14.4%; Score 93.8; DB 2; Length 51891;
Best Local Similarity 66.7%; Pred. No. 6.8e-09;
Matches 134; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 126 ACCACTGATCTGGTGGTGTGAGACACCTTACCTGATCTGTGCTGCTGCTTC 185
DB 19703 ACCATGTTGCTGGTCCCTGTTCAACACCTTTCATGAACCCCTGCTGCTGCTTCA 19644
QY 186 TGGGGCTGAGCTTCTTCAATCAAGCCGAGATCAAGAGTGTGTGACTGGAAGCG 245
DB 19643 TAGCATTCACCTTCTTCAATGAAGTCTAGAGGACAGAAATGTTGGGACTTACCGGG 19584
QY 246 CCCGCGCTTTGGCTCCAAAGCAAGTCTCAACATCTTGGCCGAGTGAAGCGCTGG 305
DB 19583 CCCAGGCTATGCTCTCCACCGCCAGTCTGAAACATCTGGGCTTGACTTTGGGCACTCC 19524
QY 306 TGGCGGCACCTGCTGCTCTGG 326
DB 19523 TCATGACCATCTGCTCATCG 19503

```

```

RESULT 12
BC009696 678 bp mRNA linear PRI 22-OCT-2001
LOCUS Homo sapiens, interferon induced transmembrane protein 2 (1-8D),
DEFINITION BC009696
ACCESSION BC009696
VERSION BC009696.1 GI:16307214
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 678)
Strausberg, R.
Direct Submission
Submitted (29-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: ggaabs-remail.nih.gov
Tissue Procurement: DCTD/DP/Genard
CNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMI)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305

```

REMARK
 COMMENT

Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL ac: <http://image.lnl.gov>

Series: IRAC Plate: 14 Row: a Column: 11
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 10835237.

Location/Qualifiers

1..678

source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="LocusID:10581"
 /db_xref="taxon:9606"
 /clone="MGC:9196 IMAGE:3876542"
 /issue_type="Lung, large cell carcinoma"
 /clone_lib="NIH MGC_68"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 31..429
 /product="interferon induced transmembrane protein 2 (1-8D)"
 /protein_id="AAH09696.1"
 /db_xref="GI:16307215"
 /translation="MNHIVQTFSPVNSGPPNYEMLEBOEVAMLGAPHPAPSTV IHRSETSVDPHVVMSLEFNTLCGLAGIAFVSKSRKRWGVDVGAQAYASTA KCLNIWALIGIFMTILVITPLVVOAQR"

CDS

BASE COUNT 146 a 231 c 148 g 153 t
 ORIGIN

Query Match 14.3%; Score 93.4; DB 9; Length 678;
 Best Local Similarity 62.8%; Pred. No. 1.8e-08;

Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

125 GACCACTGATCTGGTGGTGTTCAGACACCTCTACTGATCTGTTCCTCGGCTTC 184
 193 GACCACTGCTGGTGGTGTTCAGACACCTCTCTCTGTAACACCTGCTCGGCTTC 252
 185 CTGGCGGTGGCTCTACTCAAGGCGGAGATGAGAGGTGGTGGTGAAGGCG 244
 253 ATAGCATTCGCTCTCTCTGTAAGTCTAGAGGAGAGAGTGGTGGTGAAGGCG 312
 245 GCGCGGCTTTGGCTCTCAAGGCGGAGATGAGAGGTGGTGGTGAAGGCG 304
 313 GCCAGGCTATGCTCTCAAGGCGGAGATGAGAGGTGGTGGTGAAGGCG 372
 305 GTGCGGCT 355
 373 TTCATGACCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 423

RESULT 13

AX337050

LOCUS AX337050 905 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 7559 from Patent WO0194629.
 ACCESSION AX337050
 VERSION AX337050.1 GI:18127769

KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1 Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D. R. and Weaver, Z.
 Cancer gene determination and therapeutic screening using signature
 gene sets

JOURNAL Patent: WO 0194629-A 7559 13-DEC-2001;
 Avalon Pharmaceuticals (US)

FEATURES

Location/Qualifiers

source

1..905
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

BASE COUNT 189 a 304 c 206 g 206 t

ORIGIN

Query Match 14.3%; Score 93.4; DB 6; Length 905;
 Best Local Similarity 62.8%; Pred. No. 1.7e-08;
 Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

125 GACCACTGATCTGGTGGTGTTCAGACACCTCTACTGATCTGTTCCTCGGCTTC 184
 442 GACCACTGCTGGTGGTGTTCAGACACCTCTCTCTGTAACACCTGCTCGGCTTC 501
 185 CTGGCGGTGGCTCTACTCAAGGCGGAGATGAGAGGTGGTGGTGAAGGCG 244
 502 ATAGCATTCGCTCTCTCTCTGTAAGTCTAGAGGAGAGAGTGGTGGTGAAGGCG 561
 245 GCGCGGCTTTGGCTCTCAAGGCGGAGATGAGAGGTGGTGGTGAAGGCG 304
 562 GCCAGGCTATGCTCTCAAGGCGGAGATGAGAGGTGGTGGTGAAGGCG 621
 305 GTGCGGCT 355
 622 TTCATGACCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 672

RESULT 14

HS18D

LOCUS HS18D 905 bp mRNA linear PRI 26-MAY-1993
 DEFINITION Human 1-8D gene from interferon-inducible gene family.
 ACCESSION X57351
 VERSION X57351.1 GI:311373

KEYWORDS 1-8 gene family; 1-8D gene; interferon inducible gene.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

1 Lewin, A. R., Reid, L. B., McMahon, M., Stark, G. R. and Kerr, I. M.
 Molecular analysis of a human interferon-inducible gene family
 Eur. J. Biochem. 199 (2), 417-423 (1991)

JOURNAL

REFERENCE

AUTHORS

2 (bases 1 to 905)
 Kerr, I. M.
 Direct Submission
 Submitted (22-JAN-1991) I. M. Kerr, Imperial Cancer Research Fund,
 44 Lincoln's Inn Fields, London WC2A 3PX, U K
 On Jul 27, 1993 this sequence version replaced gi:23395.
 See X02490 for overlapping CDNA sequence.

COMMENT The Human 1-8D gene shows sequence identity to 1-8U (See X57352)
 and 9-27 (See J04164 for DNA sequence, X02491 for cDNA sequence).

FEATURES

source

1..905
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /issue_type="lymphoid"
 /clone_lib="lorisb"
 1..905
 /gene="1-8D"
 189..202
 /gene="1-8D"
 /note="ISR (interferon-stimulable response element)"
 /evidence="experimental"
 280..678
 /gene="1-8D"
 /codon_start=1
 /protein_id="CAA40625.1"
 /db_xref="GI:23396"
 /db_xref="SWISS-PROT:Q01629"

CDS

```

/translation="MNIIVQTFSPVNSGPPNYEMLKKEHEVAVLGAPHNPAPPTSTV
IHIRETSVDPHVMVSLFNTLPMNCCGFIAFAVSVKSRDRKMGDVGTGAQAVASTA
KCLINWALILIGITLILVILPILVVOAQR"
<280>.626
/gene="1-8D"
/number=1
/evidence=experimental
627..>678
/gene="1-8D"
/number=2
/evidence=experimental
698..765
/gene="1-8D"
/evidence=experimental
766..833
/rpc_type=TANDEM
/gene="1-8D"
/evidence=experimental
/rpc_type=TANDEM
886..891
/gene="1-8D"
polyA_signal
BASE COUNT 189 a 304 c 206 g 206 t
ORIGIN
Query Match 14.3%; Score 93.4; DB 9; Length 905;
Best Local Similarity 62.8%; Pred. No. 1.7e-08;
Matches 145; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 125 GACCACTTGATCTGTGGTGTTCAGACCCCTTCACTGAATCTGTCTGCTGGCTTC 184
DB 442 GACCACTGTGCTGTGTCCTGTTCAACACCTCTTCAAGAACCTGCTGCTGGCTTC 501
QY 185 CTGGCGCTGGCTCTCACTCAAGCCCGAGATCAGAAAGTGTGTGACCTGGAACG 244
DB 502 ATAGCATTCGCTTATCTCGTAGAGCTTAGGACAGAAAGATGTGTGGACGTACCGGG 561
QY 245 GCCCGCGCTTTGGCTCCAAAGCCCAAGTGTCAACAATCTGGCGCGAGTGTGACGCTG 304
DB 562 GCCCAGGCTTATGCTCCACCGCAAGTGCCTGAACATCTGGGCGCTGATTTTGGGCATC 621
QY 305 GTGGCGGCACGTGCTGCTGCTGGGCGTGTGTGACGTGTGGCTGCTGACCTG 355
DB 622 TTCATGACCATTCCTGCTGATCATCAATCCAGTGTGTGTGTCAGAGCCGAC 672

RESULT 15
BT006892 402 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens interferon induced transmembrane protein 3 (1-80)
DEFINITION mRNA, complete cds.
ACCESSION BT006892
VERSION BT006892.1 GI:30582622
KEYWORDS FLI CDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 402)
AUTHORS Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,U., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Phelan,M., and Farmer,A.
TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor
vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 402)
AUTHORS Kalline,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,
Koundinya,M., Raphael,U., Moreira,D., Kelley,T., Labaer,J., Lin,Y.,
Phelan,M., and Farmer,A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow
COMMENT This CDS clone is a part of a collection of human full length
expression clones generated by BD Biosciences Clontech and the

```

```

Harvard Institute of Proteomics. Each CDS has been cloned in two
forms: with and without stop-codon (to allow fusion with C-terminal
tag). The CDS has been directionally cloned using BD In-Fusion(TM)
cloning system between the SalI and HindIII sites of the pDNR-DUAL
vector. Additional sequences in the clone: 'ACC' after SalI site
and before 'ATG' to provide Kozak consensus sequence; 'GG' after
last codon and before HindIII site to maintain reading frame.
Clone distribution: http://bioinfo.clontech.com/orfclones.
FEATURES
Source
1..402
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00117X1.0"
/clone_id="BD Creator(TM) CDS Library derived from MGC
collection"
/lab_host="DH5alpha T1 resistant"
/note="Vector: pDNR-Dual"
1..402
/codon_start=1
/product="interferon induced transmembrane protein 3
(1-80)"
/protein_id="AAP5538.1"
/db_xref="GI:30582623"
/translation="MNIIVQTFSPVNSGPPNYEMLKKEHEVAVLGAPHNPAPPTSTV
IHIRETSVDPHVMVSLFNTLPMNCCGFIAFAVSVKSRDRKMGDVGTGAQAVASTA
KCLINWALILIGITLILVILPILVVOAQR"
ACCLINWALILIGITLILVILPILVVOAQR"
BASE COUNT 77 a 140 c 98 g 87 t
ORIGIN
Query Match 14.3%; Score 93.2; DB 9; Length 402;
Best Local Similarity 66.3%; Pred. No. 2.2e-08;
Matches 134; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 125 GACCACTTGATCTGTGGTGTTCAGACCCCTTCACTGAATCTGTGCTGGCTTC 184
DB 166 GACCACTGTGCTGTGTCCTGTTCAACACCTCTTCAAGAACCTGCTGCTGGCTTC 225
QY 185 CTGGCGCTGGCTCTCACTCAAGCCCGAGATCAGAAAGTGTGTGACCTGGAACG 244
DB 226 ATAGCATTCGCTTATCTCGTAGAGCTTAGGACAGAAAGATGTGTGGACGTACCGGG 285
QY 245 GCCCGCGCTTTGGCTCCAAAGCCCAAGTGTCAACAATCTGGCGCGAGTGTGACGCTG 304
DB 286 GCCCAGGCTTATGCTCCACCGCAAGTGCCTGAACATCTGGGCGCTGATTTTGGGCATC 345
QY 305 GTGGCGGCACGTGCTGCTGGGCGTGTGTGACGTGTGGCTGCTGACCTG 326
DB 346 CTCATGACCATTCCTGCTGATCATCAATCCAGTGTGTGTGTCAGAGCCGAC 367

Search completed: January 31, 2004, 14:30:46
Job time : 2852 secs

```